

Multivariate Elliptically Contoured Stable Distributions with Applications to BRICS Financial Data

Kimera Naradh

December, 14, 2016

Multivariate Elliptically Contoured Stable Distributions with Applications to BRICS Financial Data

by

Kimera Naradh

Dissertation submitted in fulfilment of the academic requirements for the degree of

Master of Science in Statistics

Supervisor Mr. K. Chinghamu

Co-supervisor Mr. M.J. Hammujuddy

Co-supervisor Mr. R. Chifurira



School of Mathematics, Statistics and Computer Science

Westville Campus, Durban, South Africa

December 2016



As the candidate's supervisors, we have approved this dissertation for submission

Mr. K. Chinhamu

14 December 2016
Date

Mr. M.J. Hammujuddy

14 December 2016
Date

Mr. R. Chifurira

14 December 2016
Date

Declaration - Plagiarism

I, Kimera Naradh, declare that

1. The research reported in this dissertation, except where otherwise indicated, is my original research.
2. This dissertation has not been submitted for any degree or examination at any other University.
3. This dissertation does not contain other persons' data, pictures, graphs or any other information unless specifically acknowledged as being sourced from other persons.
4. This dissertation does not contain other persons' writing, unless acknowledged as being sourced from other researchers. Where other written sources have been quoted, then
 - (a) Their words have been re-written but the general information attributed to them has been referenced, or
 - (b) Where their exact words have been used, then their writing has been placed in italics and referenced.
5. This dissertation does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being in the reference sections.

Kimera Naradh

14 December 2016

Date

Disclaimer

This document describes work undertaken as a Master's programme of study at the University of KwaZulu-Natal (UKZN). All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institution.

Abstract

Brazil, Russia, India, China and South Africa (BRICS) are regarded as the five major emerging economies where all members are a part of a select group of developing industrialized countries. In the financial industry, various models are used for the description and analysis of financial trends. One of these models is the family of stable distributions which takes into account the skewness and heavy tails that are frequent in financial data. The main objective of this study is to investigate the fit of stable distributions for exchange rates of each of the BRICS countries against the U.S. Dollar in both the univariate and multivariate cases. The data set consists of exchange rate data from the period January 2011 to January 2016.

Nolan's S_0 -parameterization stable distribution was fitted using the maximum likelihood method in the univariate case and in a fitted stable model where a GARCH (1,1) filter was applied to the returns (Stable-GARCH(1,1)). The Kolmogorov-Smirnov test and the Anderson-Darling test show that stable distributions adequately fit the returns of BRICS financial data. Value-at-Risk (VaR) calculations and VaR in-sample backtesting using the Kupiec likelihood ratio test and the Christoffersen's conditional coverage test were applied as per the International Basel Regulatory where the robustness of each model describing the financial data was evaluated. Thereafter, we proceeded to fit bivariate elliptical stable models using the Rachev-Xin-Cheng method after visualizing the scatterplot matrix of BRICS countries. This study validates the usefulness of stable distributions for modelling BRICS financial data.

‘

Keywords

BRICS, stable distributions, skewness, exchange rates, Nolan's S_0 -parameterization, stable-GARCH(1,1), Kolmogorov-Smirnov test, Anderson-Darling test, VaR, Kupiec likelihood ratio test, Christoffersen's conditional coverage test, bivariate elliptical stable model.

Dedication

This dissertation is dedicated in the loving memory of my late grandparents, Mr. Dhewlal Naradh Rughunandhan as well as Mr. Eserpersad Ramdhani and Mrs. Leela Ramdhani.

Acknowledgements

I am grateful to God for the good health and well-being necessary to complete this study. I would like to express my sincere gratitude to my supervisors Mr. K. Chinhamu as well as Mr. M.J. Hammujuddy and Mr. R. Chifurira who have been of great assistance throughout this research project with immense dedication. I am extremely thankful and indebted to them for sharing their expertise and valuable guidance to me. They have played an important role in my progress as a Master's student and have provided me with the necessary skills and knowledge needed to embark upon this research. I take this opportunity to express gratitude to Prof. John Nolan for his help and support with the STABLE package in R. I am also grateful to my parents, Mr. and Mrs. Naradh, my sister Yoshka for the unceasing love, encouragement and support. Last but not least, I gratefully acknowledge the funding received towards this Master's dissertation from the National Research Foundation (NRF). I also place on record, my sense of gratitude to one and all, who directly or indirectly, have lent their hand in this venture.

Contents

Acknowledgments	vi
1 Introduction	1
1.1 Background	1
1.2 Statement of the problem	3
1.3 Objectives of the study	4
1.4 Empirical properties of financial data	4
1.4.1 Stylized facts of financial returns	4
1.4.2 Stylized multivariate facts	6
1.5 Research layout	6
2 Literature review	7
3 Univariate stable distributions	12
3.1 Introduction to Stable distributions	12
3.1.1 Definition of stable	12
3.1.2 Alternative definitions of stability	15
3.2 Characterization and parameterization of Stable distributions	17
3.3 Distribution and density functions	22
3.4 Properties of stable laws	27
3.5 Sum of stable random variables	28
3.6 Stable parameter estimation	31
3.6.1 Univariate estimation	31
3.6.2 Maximum likelihood estimation	33

4	Multivariate stable distributions	34
4.1	Multivariate stable distributions	34
4.1.1	Multivariate stable laws	36
4.2	Discrete spectral measures	39
4.3	Multivariate stable parameter estimation	41
4.3.1	Estimating spectral measures	41
4.4	Multivariate stable densities	44
4.5	Elliptically contoured stable distributions	46
4.5.1	Elliptically contoured stable laws	46
4.5.2	Densities of elliptically contoured stable laws	47
4.5.3	Statistical analysis as elliptical stable	47
5	Volatility model: GARCH(1,1)	49
5.1	The ARCH model	49
5.2	The GARCH model	50
5.3	Parameter estimation	52
6	Methodology	55
6.1	Autocorrelation	55
6.1.1	Autocorrelation	55
6.1.2	Autocorrelation function (ACF)	55
6.1.3	Partial autocorrelation function (PACF)	56
6.1.4	Ljung-Box test for autocorrelation	56
6.2	Periodogram	57
6.2.1	Testing for hidden periodicity	58
6.3	Tests for stationarity	61
6.3.1	The unit root test	61
6.3.2	Augmentmented Dickey-Fuller test	62
6.3.3	Phillips-Perron test	63
6.3.4	Kwiatkowski-Phillips-Schmidt-Shin test	65
6.4	Measures of dependency	69
6.4.1	Scatter plot matrices	69
6.4.2	Covariance and correlation matrices	69

6.5	Goodness-of-fit tests	72
6.5.1	Kolmogorov-Smirnov test	72
6.5.2	Anderson-Darling test	73
7	Risk measures and backtesting	74
7.1	Current regulatory framework	74
7.1.1	South African implementation of Basel III Accord	75
7.2	Value-at-Risk	76
7.2.1	Steps in VaR calculations	77
7.2.2	Analyzing and interpreting VaR	78
7.3	VaR in-sample backtesting procedures	78
7.3.1	Violation ratio	79
7.3.2	Kupiec likelihood ratio test	79
7.3.3	Christoffersen conditional coverage test	80
7.3.4	VaR duration test	82
8	Analysis of BRICS financial data	85
8.1	Exchange rate data	85
8.2	Descriptive Statistics	86
8.2.1	Time series plots of daily exchange rates	86
8.3	Calculating log-returns	87
8.4	Tests for autocorrelation	90
8.4.1	Ljung-Box test for autocorrelation	91
8.5	Tests for stationarity	92
8.5.1	Checking for hidden periodicity	93
8.6	Covariance matrix	94
8.7	Correlation matrix	94
8.8	Stable parameter estimation	95
8.9	Goodness of fit tests and diagnostics	96
8.9.1	Q-Q plots of BRICS exchange rate returns	96
8.9.2	Variance stabilized P-P plots of BRICS exchange rate returns	97
8.9.3	Univariate stable density plots	99
8.10	VaR estimates and backtesting	100

8.11	Multivariate stable data analysis	102
8.11.1	Multivariate stable parameter estimation	102
9	Stable-GARCH(1,1) model	105
9.1	Stable-GARCH(1,1) model	106
9.1.1	Fitting a univariate GARCH(1,1) model to returns	106
9.1.2	Tests for autocorrelation on univariate GARCH(1,1)residuals)	110
9.2	Tests for stationarity univariate on GARCH(1,1) residuals	111
9.3	Stable parameter estimation	112
9.4	Goodness-of-fit and diagnostics	113
9.4.1	Q-Q plots for stable-GARCH(1,1) residuals	113
9.4.2	Variance stabilized P-P plots for fitted univariate stable-GARCH(1,1) residuals	114
9.4.3	stable-GARCH(1,1) density plots	116
9.5	VaR estimates and backtesting	117
9.6	Multivariate stable data analysis	119
10	Conclusion and recommendations	122
Appendix		132
	Appendix A	132
	Appendix B	136

List of Figures

8.1	Time series plots of BRICS to USD daily exchange rates	86
8.2	Time series plots of BRICS to USD exchange rate returns	87
8.3	ACF and PACF plots of daily BRICS to USD exchange rate returns	90
8.4	Q-Q plots of BRICS to USD exchange rate returns	96
8.5	[Variance stabilized P-P plots of BRICS to USD exchange rate returns . . .	97
8.6	Stable density plots of BRICS to USD daily exchange rates	99
8.7	Pairwise scatterplots of BRICS countries to USD exchange rate	102
8.8	Estimated density surface and countour plot for a bivariate elliptical stable fit of Brazillian Real and South African Rand exchange rates.	104
9.1	ACF and PACF plots of residuals for fitted univariate GARCH(1,1) model.	110
9.2	Q-Q plots for fitted univariate stable-GARCH(1,1) residuals	113
9.3	Variance stabilized P-P plots for fitted univariate stable-GARCH(1,1) residuals	114
9.4	Stable density plot of stable-GARCH(1,1) model	116
9.5	Pairwise scatterplots of univariate GARCH(1,1) residuals for BRICS countries to USD exchange rates	119
9.6	Estimated density surface and countour plot for a bivariate elliptical stable fit of Brazillian Real and South African Rand exchange rates of GARCH(1,1) residuals.	121
10.1	Estimated density surfaces and contour plots for a bivariate elliptical stable fit of BRICS countries exchange rates.	140
10.2	Bivariate elliptical stable density plots of GARCH(1,1) residuals for BRICS countries	145

List of Tables

7.1	Deviation Indicator outcomes (Christoffersen, 1998)	81
8.1	Descriptive summary statistics of daily return of BRICS to USD exchange rates.	88
8.2	Ljung-Box test of daily returns for BRICS to USD exchange rates	91
8.3	Results for ADF, PP and KPSS unit root tests for BRICS/USD Exchange rate returns	92
8.4	Fisher's Kappa test for detecting hidden periodicity in exchange rates.	93
8.5	Stable parameter estimates for daily returns under Nolan's S_0 -parameterization.	95
8.6	Goodness-of-fit tests of daily returns	98
8.7	VaR duration test	101
8.8	Bivariate elliptical stable parameter estimation for daily returns	103
9.1	ARCH LM test for heteroscedasticity	105
9.2	GARCH(1,1) parameter estimation for daily returns	106
9.3	Ljung-Box test for daily returns of the fitted GARCH(1,1) model	107
9.4	ARCH LM test for heteroscedasticity	108
9.5	Descriptive statistics of residuals to the fitted GARCH(1,1) model for daily log returns	109
9.6	Ljung-Box test for autocorrelation for univariate GARCH(1,1) residuals	111
9.7	Stationarity tests for residuals for the fitted univariate GARCH(1,1) model	111
9.8	Stable-GARCH(1,1) parameter estimates for residuals	112
9.9	Goodness-of-fit tests of Stable-GARCH(1,1) residuals	115
9.10	VaR estimates of the univariate stable-GARCH(1,1) model	117
9.11	Var backtesting for univariate stable-GARCH(1,1) model	117
9.12	VaR duration test for fitted stable-GARCH(1,1) model	118

9.13 Bivariate elliptical stable parameter estimation for GARCH(1,1) residuals . 120

Abbreviations

ACF	Autocorrelation Function
A-D	Anderson-Darling
ADF	Augmented Dickey-Fuller
ARCH	Autoregressive Conditional Heteroscedasticity
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
BRIC	Brazil, Russia, India, China
BRICS	Brazil, Russia, India, China and South Africa
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
KPSS	Kwiatkowski Phillips Schmidt Shin
K-S	Kolmogorov-Smirnov
MLE	Maximum Likelihood Estimation
PACF	Partial Autocorrelation Function
PP	Phillips Perron
Q-Q	Quantile to Quantile
VaR	Value-at-Risk

List of Special Symbols

α	index of stability
β	skewness parameter
δ	location parameter
$\stackrel{\text{d}}{=}$	equality in distribution
ϵ_t	white noise process
γ	scale parameter
$\pi(\theta)$	prior distribution
$\pi(x \theta)$	likelihood distribution
$\pi(\theta x)$	posterior distribution
\log	natural logarithm

Chapter 1

Introduction

This chapter introduces the background, statement of the research problem, objectives of the study, empirical properties of financial data and the research layout.

1.1 Background

The international competitive position of a country is considered key with regards to evaluating the success of authorities in their aim of achieving major macroeconomic goals. Global competitiveness is observed as a multidimensional phenomenon that is complex to understand using a single indicator (de Jager, 2012). Nevertheless, according to Walters & De Beer (1999), a country's real exchange rate is used to indicate the relative competitive position in international trade. Exchange rates are very important as there is an effect on a country's international relations. More specifically, imports, exports as well as foreign investment are affected by exchange rate fluctuations (Nelson, 2013). Arezki et al. (2012) indicate that increased volatility in exchange rates puts the economy in an unfavorable position through its adverse conditions on private agents consumption and investment decisions as well as commodity exporting countries experience large trade fluctuations.

The formal definition of “exchange rate” is given as the price of one currency in terms of another currency. Exchange rates are either fixed or floating. The central bank of a country decides on the fixed exchange rate whereas the floating exchange rate is decided by the market demand and supply (Picardo, 2014). Exchange rates are affected greatly

by macroeconomic triggers and are indicative of a country's financial stability. Hence, the need for reliable models that monitor the evolution of volatile exchange rates and provide necessary remedies that are useful especially in times of financial stress as the future is uncertain.

Countries in BRICS (Brazil, Russia, India, China and South Africa) are of great interest to financial analysts globally as these countries have gained prominence in the global economy given the noticeable rapid growth rate in international trading. In 2011, the Chinese city of Sanya featured South Africa for the first time in the third BRIC summit. Hence, the acronym BRICS was formed ever since. The International Monetary Fund (IMF) indicated that intra-BRICS trade is valued at billions of U.S. Dollars. Intra-BRICS trade consists of Brazil, Russia and South Africa providing the much needed natural resources for the mass industrialized needs of the Asian giants: India and China (Sule, 2011). Both BRICS and exchange rates are highly publicized news in the press and directly impact foreign investments more specifically investor psychology and confidence. In this regard, modeling BRICS exchange rates is an interesting topic of research.

Vast literature has been dedicated to modeling and evaluating changes in exchange rates and a wide variety of econometric models have been suggested by researchers. A recent study done by Caporale et al. (2016) examined the effect of macro-news on major currencies vis-à-vis the U.S. Dollar and Euro against currencies of BRICS group of countries using daily data. The estimated VAR-GARCH(1,1) model allowed for both mean and volatility spillovers as well as accounted for the impact of the global financial crisis of 2008.

Ma et al. (2013) modeled the characteristics of volatility for the exchange rate of the Chinese Yuan against the U.S. Dollar. Both the symmetric and asymmetric models of the generalized autoregressive conditional heteroscedasticity (GARCH) family were used to model the daily data. The author concluded that both models capture the characteristics of volatility in exchange rates.

In South Africa, exchange rates have been of great concern, especially in 2016, since the country's performance against major currencies (Dollar, Euro and Pound Sterling) have

weakened significantly. The Rand reached an all time low at 17.99 against the U.S. Dollar. This prompted the South African government to reshuffle cabinet ministers and reappoint a new minister of Finance (eNCA, 2016). Kemda et al. (2015) state that exchange rates, like any other financial time series, are leptokurtic and contradict the classical Gaussian assumption. Subclasses of the generalized hyperbolic distributions were compared to the Normal distribution. These authors concluded that the variance-gamma model is the most robust model for describing the South African Rand against the U.S. Dollar exchange rate at their associated VaR estimates.

The common assumption of normality for financial data tends to underestimate the probability of extreme returns, i.e. fat tails and skewness. Therefore, we can fit a stable model that takes fat tails and skewness into consideration. Stable distributions are a rich and effective class of probability laws that gives a parsimonious fit to the suggested model. This study aims to investigate the fit of the stable distribution for the exchange rates of each of the countries in BRICS using both a univariate and a multivariate time series analysis approach. The Kolmogorov-Smirnov and the Anderson-Darling goodness-of-fit tests validate the adequacy of the fitted stable models. Comprehensive VaR calculations and backtesting procedures were carried out to evaluate the robustness of each model.

We are not aware of any risk management literature relating to the application of stable, stable-GARCH(1,1) and bivariate elliptical stable case to BRICS exchange rates. Therefore, the main contribution of this study is to highlight the usefulness of stable distributions for large sets of financial data that exhibit heavy tails and skewness.

1.2 Statement of the problem

The estimation of Value-at-Risk (VaR) depends on the properties of the fitted distribution. There is a need to determine whether a distribution can adequately be used to estimate VaR of exchange rates. To fulfill this purpose, we conduct a quantitative research study in both the univariate and multivariate cases. Stable distributions, more specifically, Nolan's S_0 -parameterization combined with the GARCH(1,1) model adequately estimate VaR for BRICS exchange rate data between the time period of 2011 to 2016.

1.3 Objectives of the study

The main objective of this study is to evaluate the fit of stable distributions to the returns of BRICS financial data using both univariate and multivariate stable models.

This is achieved through:

- Time series analysis of the returns.
- Univariate stable parameter estimation using the maximum likelihood (ML) method.
- Univariate stable diagnostics: goodness-of-fit tests, such as the Kolmogorov-Smirnov test and the Anderson-Darling test.
- VaR calculations and VaR in-sample backtesting procedures using violation ratios, the Kupiec test, Christoffersen test and the Value-at-risk duration test.
- Visualizing scatterplot matrix of BRICS countries.
- Fitting appropriate multivariate stable models.
- Computing multivariate stable density plots.
- Combining stable distribution and GARCH(1,1) model. The volatility is modeled by the GARCH(1,1) process with the innovations following a stable distribution.

1.4 Empirical properties of financial data

1.4.1 Stylized facts of financial returns

The collection of observations of empirical observations and the conclusions from these observations are referred to as the stylized facts of financial returns. These apply to most daily series of risk factor changes, for example, log-returns on equities, exchange rates and commodity prices. These observations are deeply associated with econometrics that they are now considered facts in their own rights. We list a detailed version of the stylized facts below:

- (i) Returns are not independent and identically distributed but they exhibit serial correlation.

- (ii) Squared returns series show profound serial correlation.
- (iii) Conditional expected returns are close to zero.
- (iv) Time-varying volatility.
- (v) Returns are leptokurtic or heavy-tailed.
- (vi) Extreme returns are clustered.

In this study, we focus mainly on the properties of volatility clustering, non-Normality, heavy tails and longer interval return series.

Volatility clustering

Volatility clustering is the tendency for extreme returns to be followed by other extreme returns. Volatility is modelled as conditional standard deviation of financial returns and, even though conditional expected returns are close to zero, the presence of volatility clustering shows that conditional standard deviations are changing in a predictable way.

Non-Normality and heavy tails

The Normal distribution is a poor model for daily returns and the Jarque-Bera test (based on empirical skewness and kurtosis) may, in some cases, reject the assumption of Normality. Daily financial returns have a higher kurtosis and are said to be leptokurtic. That is, there is a narrow center with longer heavier tails than the Normal distribution

Longer interval return series

As the interval of the returns is increased from daily to weekly, monthly, quarterly and yearly data, volatility clustering becomes less pronounced and returns are less heavy-tailed and i.i.d. If we have a sample with n returns measured over some time interval, for example daily or weekly, and if we aggregate these to form longer-interval logarithmic returns then the k period log-return at time t is given by:

$$Y_t^k = \ln\left(\frac{S_t}{S_{t-k}}\right) = \ln\left(\frac{S_t}{S_{t-1}} \cdots \frac{S_{t-k+1}}{S_{t-k}}\right) = \sum_{j=0}^{k-1} Y_{t-j} \quad (1.1)$$

We can form a sample of non-overlapping k period returns $\{Y_t^k : t = k, 2k, \dots, [\frac{n}{k}k]\}$. A central limit effect occurs due to the sum structure of the k -period returns. The distribution becomes less leptokurtic and more Normal as k increases. The central limit theorem applies to many stationary time series processes, including GARCH models.

1.4.2 Stylized multivariate facts

Financial risk managers are seldom interested in one time series but rather a multiple series of financial risk factors. Consider the multivariate return data $\mathbf{Y}_1, \dots, \mathbf{Y}_n$. Each component series $Y_{1,j}, \dots, Y_{n,j}$ for $j = 1, \dots, d$ is a series formed by log-differencing daily commodity prices or exchange rates. Consider the following stylized facts:

- (i) There is little evidence of cross-correlation for multivariate returns with an exception for contemporaneous returns.
- (ii) There is, however, profound evidence of cross-correlation for a multivariate series of absolute returns.
- (iii) Contemporaneous returns (correlations) between series vary over time.
- (iv) Extreme returns in one series coincide with extreme returns in other series (McNeil et al., 2005).

1.5 Research layout

This dissertation consists of ten chapters. Subsequent to this introductory chapter, the second chapter provides a literature review. Chapter 3 introduces univariate stable distributions with several definitions and various properties. In Chapter 4, we discuss the theory of multivariate stable distributions. Chapter 5 describes the GARCH(1,1) model and Chapter 6 provides research methodology. Chapter 7 discusses risk measures and backtesting procedures, thereafter, Chapter 8 presents the data analysis of BRICS exchange rate returns. Chapter 9 combines the stable and GARCH(1,1) model. Finally, Chapter 10 summarizes the findings and concludes this study.

Chapter 2

Literature review

In this section, we overview research done on similar topics.

Nolan (2003) investigated the use of stable distributions with financial data based on the British Pound versus the German Mark exchange rate. The data consisted of daily exchange rate for the period 2 January 1990 to 21 May 1996. The returns for the data set were computed and parameter estimation was carried out using the maximum likelihood (ML) method. The data were analyzed by the fitted stable model and the suggested Normal fitted model. Nolan (2003) also studied the monthly exchange rates between the U.S. Dollar and the Tanzanian Shilling. The data ranged from January 1975 to September 1997. The returns of the data were computed and the parameter estimates were found using the ML method. The study found that the Tanzanian Shilling exchange rate was subject to more extreme fluctuations.

McCulloch (1997) investigated the suitability of stable distributions using data from the stock market namely the stock price data known as the Centre for Research in Security Prices (CRSP). This data set was analyzed over forty years from January 1953 to December 1992. The ML estimates were calculated as well as the quantile estimates. The goodness-of-fit was studied using graphical methods by observing the P-P plot and the stable density plot. Diagnostics showed a close fit.

Nolan (2003) investigated the joint distribution between the German Mark and the Japanese Yen. The main interest in both currencies is to see if the joint distribution is bivariate stable and in estimating the fit. A sequence of smoothed Q-Q plots and variance-stabilized P-P plots were projected in 8 different directions (a restriction is placed on the right half-plane as the left half-plane is merely a reflection of the right half-plane). This multivariate study was adequate in describing the data except on the extreme tails. Projection functions $\alpha(t)$, $\beta(t)$, $\gamma(t)$ and $\delta(t)$ based on stable distributions were estimated and were also used to formulate an estimate of the spectral measure using the projection method. Four plots of each of the parameter estimates were analyzed. The plots for the skewness $\beta(t)$ and the scale function $\gamma(t)$ are computed from the estimated spectral measure. The curves are identical to the direct, separate estimates of the directional parameters. The fitted spectral measure was used to plot the bivariate density.

Press (1972) studied an application on portfolio analysis. The returns of the price per unit asset were described. In the application, the returns followed a univariate symmetric stable distribution. Portfolio management firms would most likely be interested in the joint return behavior of many such portfolios. It was found that the vector of portfolio returns will have jointly stable components.

Chinhamu et al. (2015) investigated the best generalized hyperbolic distribution to fit gold price returns where comparisons to fitted stable distributions were made. The adequacy of the distributions were assessed by the Anderson-Darling test, Bayesian information criterion, Akaike information criterion and backtesting of VaR estimates. It was found that the best model for gold returns differ at different VaR levels and that the stable distribution along with the generalized hyperbolic distribution favorably describe extreme risk in gold returns.

Mandalos (2014) investigated the relationship between the South African Rand and the U.S. Dollar exchange rate including the macroeconomic changes between the two countries for further interpretation and prediction of the exchange rate in the future. The study aimed to find the determinants of the nominal exchange rates over the period after South Africa's financial liberalization starting in 2002. A time series analysis was carried out

using an empirical model linked to theoretical determinants to exchange rate and was used to provide long-run and short-run effects on exchange rates. Mandalos (2014) also noted that there is an absence of studies in exchange rate movements in developing countries such as South Africa.

Campa et al. (2002) used currency data from the BMF, the Commodities and Futures exchange in Sao Paulo, Brazil, to investigate the market expectations of the Brazilian Real to the U.S. Dollar exchange rate over the period from 1994 to 1997. Probability density functions were derived to analyze the expected future exchange rates and investigate the credibility governing regimes on the exchange rate, namely the “crawling peg” and target zone (“Maxiband”) regime. The analysis is based on the risk-neutral probability density function. The study concluded that the credibility of the target zone was poor prior to February 1996 and improved thereafter.

Ločmelis et al. (2015) analyzed the impact on the changes in dynamic linkages between the Russian, U.S. and EU stock markets amidst Russia’s financial crisis that started in 2014. A structural break analysis was performed to identify a possible period of tranquility in the Russian stock market and a date at which the financial crisis period started. Thereafter, cointegration, Granger-causality, impulse response, variance decomposition and GARCH-BEKK tests were conducted to draw comparisons between the long-run and short-run volatility, and shock spillover linkages during the financial crisis and stable periods.

Murari & Sharma (2013) investigated the dynamics of the Indian Rupee fluctuations against the U.S. Dollar using observations from 2001 to 2013. Ordinary least squares (OLS) modeling was carried out on the log-computed variables to investigate the determinants of the Rupee fluctuations against the U.S. Dollar. Six factors were found to be behind the fluctuation and were modeled by multivariate regression analysis. Modeling exchange rates through various econometric techniques based on currency rates remain an area for further research.

Dasgupta (2014) studied the interrelationships, interdependence, integration and dynamic

linkages between the “BRIC” countries with reference to India. Data from the daily closing values of the BRIC indices were used. Tests such as the Jarque-Bera test, the Augmented Dickey-Fuller and Phillips-Perron tests for identifying Normality and stationarity were carried out. The study also used the Johansen-Juselius and the Engle-Granger cointegration tests as well as the Granger causality tests to investigate the short-run and long-run integration, and interrelationships of the BRIC stock market. The study was made more reliable by the use of vector autoregression and variance decomposition analysis. It was found that the Indian stock market had a strong impact on the Brazilian and Russian stock markets and, in general, the study also found that the BRIC stock markets were attractive to global investors and emphasized the dominance of the Indian stock market among the “BRIC” countries.

Ijumba (2013) investigated levels of independence and dynamic linkages among the BRICS countries using vector autoregressive, univariate GARCH(1,1) and multivariate GARCH(1,1). The data consisted of the weekly returns from January 2000 to December 2012. Results from the VAR model showed unidirectional linear dependencies of the Chinese and Indian markets on the Brazilian stock market. The univariate GARCH model implied that the stock returns of China seemed to be most volatile followed by Russia whereas the South African stock market was found to be the least volatile. Results obtained from the multivariate GARCH model yield similar conclusions. This study illustrated that interdependence amongst BRICS countries cannot be rejected and multiple factors, besides internal markets, may affect correlation and volatility among BRICS countries.

Pradhan et al. (2013) examined economic growth and financial development using panel data vector autoregression. The study found bidirectional causality between economic growth and financial development, and highlighted the importance of economic policies to acknowledge financial growth in emerging BRICS economies.

Nolan (2014) examined a small portfolio example with three assets, namely: Ford, IBM, Procter and Gamble. The closing prices were obtained for ten years. In the univariate example, changing volatility was evident mostly in 2008. A GARCH(1,1) filter was applied to the data. The pairwise scatterplots of the residuals displayed a roughly elliptical pattern

and this implied the estimation of a jointly stable three-dimensional elliptical model for the data.

To the best of our knowledge, there are limited studies on using stable distributions to fit BRICS financial data, more specifically exchange rates. Therefore, this is the only study that acknowledges modeling exchange rates among BRICS countries using univariate and multivariate stable analyses.

Chapter 3

Univariate stable distributions

This chapter provides an introduction to univariate stable distributions with several definitions, theorems and properties.

3.1 Introduction to Stable distributions

Stable distributions are a four-parameter family of models that generalize the normal model. Models that are based on stable laws and properties ideally describes real data well over its range, provides robust models for compounding returns as well as account for heavy tails and skewness. With the progression of statistical software, the practical use of stable distributions is advocated in finance. While there are many other classes of models that may provide a good fit for financial data sets, however, they lack the favorable features mentioned earlier (Nolan, 2014).

The theory of stable distributions stems from the pioneering work of Paul Lévy in the 1930s.

3.1.1 Definition of stable

Definition 3.1.1 (Nolan, 2003)

- (i) The sum of two Normally distributed random variables yields a Normal random variable. If Y is Normal, then Y_1 and Y_2 are independent and identical to Y with

any positive constants a and b .

$$aY_1 + bY_2 \stackrel{d}{=} cY + d \quad (3.1)$$

for $c \geq 0$ and $d \in \mathbb{R}$ where $\stackrel{d}{=}$ denotes equality in distribution.

(ii) Any random variable is *symmetrically stable* if it is stable and symmetrically distributed around 0, that is, $Y \stackrel{d}{=} -Y$.

(iii) A random variable is *strictly stable* if $d = 0$.

The addition rule for independent Normally distributed random variables states that the mean of the sum is the sum of the means and the variance is the sum of the variances. Suppose $Y \sim N(\mu, \sigma^2)$, terms on the left-hand-side of equation (3.1) are $N(a\mu, (a\sigma^2))$ and $N(b\mu, (b\sigma^2))$. On the right-hand-side, we have $N(c\mu + d, (c\sigma^2))$. From the addition rule, we have: $c^2 = a^2 + b^2$ and $d = (a + b - c)\mu$. The use of the word *stable* is justified because the shape does not change under addition as indicated by equation (3.1). In the literature, many authors coin the phrase *sum stable* to further emphasize the fact that equation (3.1) deals with addition. The term *stable* should also be distinguished between these distributions described by max-stable, min-stable, multiplication stable and geometric stable distributions. Different terms were used in older literature. *Stable* was referred to as what is now *strictly stable*, the term *quasi-stable* was used to as what we now refer to as *stable*. Two random variables Y and Z are said to be of a similar type if there exists constants $A > 0$ and $B \in \mathbb{R}$ such that $Y \stackrel{d}{=} AZ + B$. Then, the definition of *stable* can be restated as $aY_1 + bY_2$ and of similar type as Y .

Stable distributions are attractive in theory but they are difficult to implement. There exist three special cases which can be expressed in closed-form densities. It can be verified that they are stable in nature. The family of alpha-stable distributions is a rich class and includes the Normal, Cauchy and Lévy distributions as subclasses, which are described below by their density functions. The stable parameters α, β, γ and δ are defined in more detail in Section 3.2.

(i) Normal/Gaussian distribution $Y \sim N(\mu, \sigma^2)$

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(y - \mu)^2\right], \quad -\infty < y < \infty \quad (3.2)$$

The Normal distribution is stable with $\alpha = 2$ and skewness parameter $\beta = 0$.

(ii) Cauchy distribution $Y \sim \text{Cauchy}(\gamma, \delta)$

$$f(y) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (y - \delta)^2}, \quad -\infty < y < \infty \quad (3.3)$$

The Cauchy laws are stable with $\alpha = 1$ and $\beta = 0$.

(iii) Lévy distribution $Y \sim \text{Lévy}(\gamma, \delta)$

$$f(y) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(y - \delta)^{3/2}} \exp\left[-\frac{\gamma}{2(y - \delta)}\right], \quad \delta < y < \infty \quad (3.4)$$

The Lévy distributions are stable with $\alpha = \frac{1}{2}$ and $\beta = 1$.

The Normal and Cauchy distributions are both symmetric and bell-shaped curves. The main difference between the two is that the Cauchy distribution has heavier tails. However, the Lévy distribution is skewed and has heavier tails than the Cauchy distribution (Nolan, 2015).

The Normal distribution is widely used in financial modeling partly because of its favorable analytical properties, which are also shared by other members of the stable distribution family (Yang, 2012a).

The reasons why the Normal distribution is popular in financial modeling are:

- It is a relatively straightforward and practical distribution where numerical methods can be implemented.
- The Central Limit Theorem and the Law of Large Numbers are properties that simplify problems in Statistics by working with distributions that are approximately Normal.
- Normally distributed random variables assume values around the central mean where the odds of deviation from the mean exponentially decrease as one moves

away from it.

Well-known financial frameworks based on the Normal distribution are (Stoyanov et al., 2011):

- (a) Black-Scholes option pricing model.
- (b) Capital asset pricing model.
- (c) Markowitz's modern portfolio theory.

3.1.2 Alternative definitions of stability

Definition 3.1.2 (Nolan, 2003)

Non-degenerate Y is stable if and only if $\forall n > 1, \exists$ constants $c_n > 0$ and $d_n \in \mathbb{R}$ such that

$$Y_1 + \dots + Y_n \stackrel{d}{=} c_n Y + d_n \quad (3.5)$$

where Y_1, \dots, Y_n are independent and identical copies of Y and are strictly stable if $d_n = 0 \forall n$. The constant c_n must be of the form $c_n = n^{1/\alpha}$ for some $\alpha \in (0, 2]$. Distributional properties of Y are used in both definitions above. Another distributional characterization is given by the Generalized Central Limit Theorem. (See Appendix A). The most accurate way to describe stable distributions is by a characteristic function or Fourier transform. For a random variable Y with distribution function $F(y)$, the characteristic function is defined as

$$\phi(t) = E(e^{itY}) = \int_{-\infty}^{\infty} e^{itY} dF(y) \quad (3.6)$$

where $\phi(t)$ determines the distribution of Y and the sign function is defined as:

$$\text{sign } y = \begin{cases} -1, & y < 0; \\ 0, & y = 0; \\ 1, & y > 0. \end{cases}$$

Definition 3.1.3 (Nguyen & Sampson, 1991)

A distribution function $F(y)$ is said to be univariate stable if for every $b_1 > 0, b_2 > 0$, real c_1, c_2 , there is a corresponding positive number b and a real number c such that for every

scalar y , where $-\infty < y < \infty$,

$$F\left(\frac{y - c_1}{b_1}\right) * F\left(\frac{y - c_2}{b_2}\right) = F\left(\frac{y - c}{b}\right) \quad (3.7)$$

where $*$ denotes the convolution operator.

A univariate stable distribution has a characteristic function ϕ given by

$$\phi(t) = i\mu t - \gamma |t|^\alpha \left[1 + i\beta \frac{t}{|t|} \omega(t, \alpha) \right] \quad (3.8)$$

where $-\infty < t < \infty$, with given $-\infty < \mu < \infty$, $-1 \leq \beta \leq 1$, $0 < \alpha < 2$, $\frac{t}{|t|}$ at $t = 0$ and for all t :

$$\omega(t, \alpha) = \begin{cases} \tan \frac{\pi\alpha}{2}, & \alpha \neq 1; \\ \frac{2}{\pi} \ln(|t|), & \alpha = 1. \end{cases}$$

A random variable with a stable distribution can be characterized by the identical distribution to that of a random variable and a linear combination of n independant copies of that random variable. It also depends on the interrelationships of the coefficients of the linear form.

Definition 3.1.4 (Nolan, 2015)

A random variable Y is *stable* if and only if $Y \stackrel{d}{=} aZ + b$, where $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $a \neq 0$, $b \in \mathbb{R}$ and Z is a random variable with characteristic function

$$E(e^{itZ}) = \begin{cases} \exp(-|t|^\alpha [1 - i\beta \tan \frac{\pi\alpha}{2} (\text{sign}(t))]), & \alpha \neq 1; \\ \exp(-|t| [1 + i\beta \frac{2}{\pi} (\text{sign}(t)) \log |t|]), & \alpha = 1. \end{cases} \quad (3.9)$$

The distributions are symmetric when $\beta = 0$ and $b = 0$. Then, the characteristic function of aZ has the form $\phi(t) = e^{-a^\alpha |t|^\alpha}$.

3.2 Characterization and parameterization of Stable distributions

Stable distributions are described by four parameters, namely α, β, γ , and δ . The index of stability/index of law/characteristic of exponent is explained by the parameter α where $0 < \alpha < 2$. Skewness is denoted by the parameter β where $-1 < \beta < 1$. The distribution is symmetric if $\beta = 0$. If $\beta > 0$, the distribution is skewed to the right and if $\beta < 0$, then the distribution is skewed to the left. The shape of the distribution is determined by α and β . The scale parameter is denoted by $\gamma > 0$. The parameter δ denotes the rightward or leftward shift of the distribution. It is called the location parameter. The distribution has a leftward shift if $\delta < 0$. Conversely, the distribution has a rightward shift if $\delta > 0$.

Multiple parameterizations are used to describe stable laws. This is due to a historical evolution and the many problems that have been observed when analyzing stable distributions. If one works with fitting data, or numerical work, then the first parameterization is preferred. Yet, if there is a desire to work with simple algebraic structures, then another parameterization is advised, and if one studies the analytical properties of strictly stable laws, then another parameterization would be useful. The notation $S(\alpha, \beta, \gamma, \delta, k)$ is used to describe the class of stable laws. The four parameters α, β, γ and δ are unknown and need to be estimated. The integer k distinguishes between the different parameterizations (Nolan, 2015).

Definition 3.2.1 (Nolan, 2015)

Nolan's S_0 -parametrization A random variable Y is $S(\alpha, \beta, \gamma, \delta; 0)$ if

$$Y \stackrel{d}{=} \begin{cases} \gamma(Z - \beta \tan \frac{\pi\alpha}{2}) + \delta, & \alpha \neq 1; \\ \gamma Z + \delta, & \alpha = 1. \end{cases} \quad (3.10)$$

where $Z \equiv Z(\alpha, \beta)$ has characteristic function (3.9). In this case Y has characteristic function:

$$E(e^{itY}) = \begin{cases} \exp(-\gamma^\alpha |t|^\alpha [1 + i\beta (\tan \frac{\pi\alpha}{2}) (\text{sign}(t)) \times (|\gamma t|^{1-\alpha} - 1)] + i\delta t), & \alpha \neq 1; \\ \exp(-\gamma |t| [1 + i\beta \frac{2}{\pi} (\text{sign}(t)) \times \log(\gamma |t|)] + i\delta t), & \alpha = 1. \end{cases} \quad (3.11)$$

Nolan (2014) recommends using the S_0 -parameterization for statistical inferences, and numerical purposes, as it has the simplest form for the characteristic function that is continuous in all four parameters. The S_0 -parameterization acknowledges a location-scale family. If $Z \sim S(\alpha, \beta, \gamma, \delta; 0)$, then for $\alpha \neq 0$, $b \in \mathbb{R}$, $aZ + b \sim S(\alpha, \text{sign}(\alpha)\beta, |a|\gamma, a\delta + b; 0)$.

Definition 3.2.2 Nolan's S_1 -parametrization (Nolan, 2015)

A random variable Y is $S(\alpha, \beta, \gamma, \delta; 1)$ if

$$Y \stackrel{d}{=} \begin{cases} \gamma Z + \delta & \alpha \neq 1; \\ \gamma Z + (\delta + \beta \frac{2}{\pi} \gamma \log \gamma) & \alpha = 1, \end{cases} \quad (3.12)$$

where $Z \equiv Z(\alpha, \beta)$ has characteristic function (3.9). In this case, Y has characteristic function:

$$E(e^{itY}) = \begin{cases} \exp(-\gamma^\alpha |t|^\alpha [1 - i\beta (\tan \frac{\pi\alpha}{2}) (\text{sign}(t))] + i\delta t) & \alpha \neq 1; \\ \exp(-\gamma |t| [1 + i\beta \frac{2}{\pi} (\text{sign}(t)) \log(\gamma |t|)] + i\delta t), & \alpha = 1. \end{cases} \quad (3.13)$$

Yang (2012a) describes the following Zolotrev's parameterizations.

Definition 3.2.3 Zolotrev A-parameterization

A random variable Y is $S(\alpha, \beta, \gamma, \delta; A)$ if the characteristic function can be described as follows:

$$E(e^{itY}) = \begin{cases} \exp(\gamma[it\delta - |t|^\alpha + it|t|^{\alpha-1}\beta\tan\frac{\pi\alpha}{2}]) & \alpha \neq 1; \\ \exp(\gamma[it\delta - |t|^\alpha - i\beta\frac{2}{\pi}t\log|t|]), & \alpha = 1. \end{cases} \quad (3.14)$$

The characteristic functions in (3.14) are discontinuous in the parameters determining them. Discontinuities exist at all points of the form $\alpha = 1$ and $\beta \neq 0$. If we take the limits $\alpha^* \rightarrow 1$ ($\alpha^* \neq 1$), $\beta^* \rightarrow \beta \neq 0$, $\gamma^* \rightarrow \gamma$ and $\delta^* \rightarrow \delta$ it does not yield a stable law with parameters $\alpha = 1, \beta, \gamma$ and δ but more especially it does not yield an appropriate distribution in the limit. The entire measure tends to infinity. By adding a shift to the location parameter, $-\beta\tan\frac{\pi\alpha}{2}$, the discontinuity is removed.

Definition 3.2.4 Zolotrev M-parameterization

A random variable Y is $S(\alpha, \beta, \gamma, \delta; M)$ if the characteristic function can be described as follows:

$$E(e^{itY}) = \begin{cases} \exp(\gamma[it\delta - |t|^\alpha + it(|t|^{\alpha-1} - 1)\beta\tan\frac{\pi\alpha}{2}]) & \alpha \neq 1; \\ \exp(\gamma[it\delta - |t|^\alpha - i\beta\frac{2}{\pi}t\log|t|]), & \alpha = 1. \end{cases} \quad (3.15)$$

One should note the similarities between Nolan's S_0 -parameterization and Zolotrev M-parameterization where changes only in γ and δ are made so that they are more accommodating to the classical sense of the scale and location parameters. Likewise, the same relationship applies to Nolan S_1 -parameterization and Zolotrev A-parameterization. The cumulative distribution function satisfies $F(y; \gamma) = F(y/\gamma; 1)$, where γ is the scale parameter in the classical definition, for some distributions. Nolan's S_0 -parameterization, S_1 -parameterization and the scale parameter γ belongs to this category. Some parameterizations mimic the scale parameter, that is, we observe a combination of scale parameters and some other parameters such as Zolotrev A-parameterization.

Definition 3.2.5 Zolotrev B-parameterization

A random variable Y is $S(\alpha, \beta, \gamma, \delta; B)$ if the characteristic function can be described as follows:

$$E(e^{itY}) = \begin{cases} \exp(\gamma[it\delta - |t|^\alpha \exp(-i\frac{\pi}{2}\beta K(\alpha)\text{sign}(t))]) & \alpha \neq 1; \\ \exp(\gamma[it\delta - |t|^\alpha (\frac{\pi}{2} + i\beta \log |t|\text{sign}(t))]), & \alpha = 1, \end{cases} \quad (3.16)$$

where $K = \alpha - 1 + \text{sign}(1 - \alpha)$, the parameters have the same domain of variation as in the A-parameterization. The B-parameterization as in the A-parameterization show that stable law are discontinuous at points of the form $\alpha = 1$. Nevertheless, the B-parameterization has a limit distribution that exists and is stable in its distribution as $\alpha^* \rightarrow 1_+, \beta^* \rightarrow \beta, \gamma^* \rightarrow \gamma$ and $\delta^* \rightarrow \delta$. $\rightarrow 1_+$ denotes the convergence to 1 from above.

Zolotarev (1986) describes the following parameterizations:

Definition 3.2.6 Zolotrev C-parameterization

A random variable Y is $S(\alpha, \beta, \gamma, \delta; C)$ if the characteristic function can be described as follows:

$$E(e^{itY}) = -\delta|t|^\alpha \exp\left(-i\left(\frac{\pi}{2}\right)\theta_\alpha \text{sign } t\right) \quad (3.17)$$

where the parameters vary within their limits: $0 < \alpha \leq 2, \delta > 0, |\theta| \leq \theta_\alpha = \min(1, 2/\alpha - 1)$.

Definition 3.2.7 Zolotrev E-parameterization

A random variable Y is $S(\alpha, \beta, \gamma, \delta; E)$ if the characteristic function can be described as follows:

$$E(e^{itY}) = -\nu^{\frac{1}{2}} \left(\log|t| + \tau - i\left(\frac{\pi}{2}\right)\theta \text{sign } t \right) + \mathbb{C}(\nu^{-\frac{1}{2}} - 1) \quad (3.18)$$

where $\mathbb{C} \approx 0.577$ (Euler constant) and the parameters vary within their limits: $\nu \geq \frac{1}{4}, |\theta| \leq (1, 2\sqrt{\nu} - 1), |\tau| < \infty$.

For stable distributions, it is vital to determine the parameterization before random variable generation, hypothesis testing and parameter estimation. Some conversions between the parameterizations are listed below.

$$S_0 \rightarrow S_1$$

$$\beta_1 = \beta_0, \gamma_1 = \gamma_0, \delta_1 = \begin{cases} \delta_0 - \beta_0 \gamma_0 \tan \frac{\pi \alpha}{2} & \alpha \neq 1, \\ \delta_0 - \beta_0^2 \gamma_0 \ln \gamma_0 & \alpha = 1. \end{cases}$$

$$(M) \rightarrow (A)$$

$$\begin{aligned} \beta_A &= \beta_M, \delta_A = \delta_M - \beta_M \tan \frac{\pi \alpha}{2}, \gamma_A = \gamma_M & \text{if } \alpha \neq 1; \\ \beta_A &= \beta_M, \delta_A = \delta_M, \gamma_A = \gamma_M & \text{if } \alpha = 1. \end{aligned}$$

Further conversions are described in Appendix A.

3.3 Distribution and density functions

Yang (2012a) suggests that stable distributions do not have explicit closed-form cumulative distribution functions or probability density functions. Nolan, 2001 explores a numerical method for computing densities and Zolotarev (1986) shows in detail the integral form of density functions for stable distributions.

Zolotarev (1986) states the integral formula in the M-parameterization, defined by

$$\zeta = \zeta(\alpha, \beta) = \begin{cases} -\beta \tan \frac{\pi\alpha}{2} & \alpha \neq 1; \\ 0 & \alpha = 1. \end{cases} \quad (3.19)$$

$$\theta_0 = \theta_0(\alpha, \beta) = \begin{cases} -\frac{1}{\alpha} \arctan(\beta \tan \frac{\pi\alpha}{2}) & \alpha \neq 1; \\ \frac{\pi}{2} & \alpha = 1. \end{cases} \quad (3.20)$$

$$c_1(\alpha, \beta) = \begin{cases} \frac{1}{\pi} \left(\frac{\pi}{2} - \theta_0 \right) & \alpha < 1; \\ 0 & \alpha = 1; \\ 1 & \alpha > 1. \end{cases} \quad (3.21)$$

$$V(\theta; \alpha, \beta) = \begin{cases} (\cos \alpha \theta_0)^{1/\alpha-1} \left(\frac{\cos \theta}{\sin \alpha(\theta_0 + \theta)} \right)^{\alpha/\alpha-1} \frac{\cos(\alpha \theta_0 + (\alpha-1)\theta)}{\cos \theta} & \alpha \neq 1; \\ \frac{2}{\pi} \left(\frac{\pi/2 + \beta \theta}{\cos \theta} \right) \exp \left(\frac{1}{\beta} \left(\frac{\pi}{2} + \beta \theta \right) \tan \theta \right) & \alpha = 1, \beta \neq 0. \end{cases} \quad (3.22)$$

The integral formula is very complex and is the reason as to why there exists various setbacks for the applications of stable distributions.

Theorem 3.3.1 (Nolan, 2015)

All non-degenerate stable distributions are continuous distributions with an infinitely differentiable density where $f(x|\alpha, \beta, \gamma, \delta; k)$ denotes the density function and $F(x|\alpha, \beta, \gamma, \delta; k)$ denotes the distribution function of an $S(\alpha, \beta, \gamma, \delta; k)$ distribution. When the scale parameter $\gamma = 1$ and the location parameter $\delta = 0$, the distribution is standardized. The density function and distribution function of the standardized distribution are denoted by $f(x|\alpha, \beta; k)$ and $F(x|\alpha, \beta; k)$, respectively. Stable densities are supported by the entire real line or half a line. The half-line situation occurs when $\alpha < 1$ and $\beta = -1$ or $\beta = 1$. More detailed limits are given by Lemma 3.3.1.

Lemma 3.3.1 (Nolan, 2015)

$$\text{support} f(y|\alpha, \beta, \gamma, \delta; 0) = \begin{cases} [\delta - \gamma \tan \frac{\pi\alpha}{2}, \infty) & \alpha < 1 \text{ and } \beta = 1 \\ (-\infty, \delta + \gamma \tan \frac{\pi\alpha}{2}] & \alpha < 1 \text{ and } \beta = 1 \\ (-\infty, +\infty) & \text{otherwise.} \end{cases} \quad (3.23)$$

$$\text{support} f(y|\alpha, \beta, \gamma, \delta; 1) = \begin{cases} [\delta, \infty) & \alpha < 1 \text{ and } \beta = 1 \\ (-\infty, \delta] & \alpha < 1 \text{ and } \beta = 1 \\ (-\infty, +\infty) & \text{otherwise.} \end{cases} \quad (3.24)$$

The term $\tan \frac{\pi\alpha}{2}$ is a constant as is seen often when working with stable distributions. We observe as $\alpha \uparrow 1$, then $\tan \frac{\pi\alpha}{2} \downarrow -\infty$. There is a discontinuity at $\alpha = 1$. This is troublesome when working with stable distributions. It is also possible that if $|\beta| = 1$, then as $\alpha \uparrow 1$, the support in Lemma 3.3.1 tends to \mathbb{R} naturally.

The reflection property is a basic fact of stable distributions.

Property 3.3.1 Reflection Property (Nolan, 2015)

For any α and β , $P \sim S(\alpha, \beta; k)$ where $k = 0, 1, 2$

$$P(\alpha, -\beta) \stackrel{d}{=} -P(\alpha, \beta) \quad (3.25)$$

The random variable $P(\alpha, \beta)$ have density and distribution functions that satisfies: $f(y|\alpha, \beta; k) = f(-y|\alpha, -\beta; k)$ and $F(y|\alpha, \beta; k) = 1 - F(-y|\alpha, -\beta; k)$. If $Y \sim S(\alpha, \beta, \gamma, \delta; k)$ then $-Y \sim S(\alpha, -\beta, \gamma, -\delta; k)$. Therefore, $f(y|\alpha, \beta, \gamma, \delta; k) = f(-y|\alpha, -\beta, \gamma, -\delta; k)$ and $F(y|\alpha, \beta, \gamma, \delta; k) = 1 - F(-y|\alpha, -\beta, \gamma, -\delta; k)$.

When $\beta = 0$, the reflection property suggests $f(y|\alpha, 0; k) = f(-y|\alpha, 0; k)$. This implies that the density and distribution functions are symmetric around 0. We may graphically assess and observe that as α decreases the peaks of bell-shaped symmetric stable distributions get higher and the region closest to the peak gets lower and the tails get heavier. The distribution is skewed with the right tail heavier than the left tail $P(Y > y) > P(Y < -y)$ for large $y > 0$ when $\beta > 0$. A stable distribution is considered totally skewed to the right when $\beta = 1$. $\beta < 0$ is a reflection of $\beta > 0$ by the reflection property. Here, the left tail is heavier than the right tail. A stable distribution is considered to be *totally skewed* to the left when $\beta = -1$. We have a non-standardized Normal distribution when $\alpha = 2$. In this case, $\tan \frac{\pi\alpha}{2}$ in equation (3.9). The distribution is always symmetric and the characteristic function is always real regardless of the value of β . Symbolically, it can be represented as $P(2, -\beta) \stackrel{d}{=} -P(2, \beta)$. Generally, all stable distributions get closer to being symmetric as $\alpha = 2$ and β is difficult to estimate precisely which makes it insignificant in applications.

Stable distributions do not have a known formula for the location of the mode. All stable distributions can be described as unimodal. $m(\alpha, \beta)$ denotes the mode of $Z \sim S(\alpha, \beta; 0)$ distribution. $m(\alpha, -\beta) = -m(\alpha, \beta)$, by the reflection property. It can also be numerically observed that $P(Z > m(\alpha, \beta)) > P(Z < m(\alpha, \beta))$ when $\beta > 0$ (more mass to the right of the mode). By the reflection property, when $\beta < 0$, then $P(Z > m(\alpha, \beta)) < P(Z < m(\alpha, \beta))$ and there is more mass to left of the mode. If $\beta = 0$, then $P(Z > m(\alpha, \beta)) = P(Z < m(\alpha, \beta)) = 1/2$ (Nolan, 2015).

[**Note:** The above statements are all in the Nolan's S_0 -parameterization only]

Property 3.3.2 (Yang, 2012a)

Let $Y \sim S(\alpha, \beta, \gamma, \delta)$ and $f(y)$ and $F(y)$, be its density and distribution function, respectively. When $\alpha = 2$, the Normal distribution has asymptotic tail properties and when $\alpha < 1$, stable distributions have one tail when $\alpha < 1$ and $\beta = \pm 1$, and both tails otherwise, where there are cases that are asymptotic power laws with heavy tails.

(i) **Paretian tail density**

Both tail densities and probabilities of non-Normal stable distributions are asymptotically power laws. If $0 < \alpha < 2$ and $-1 < \beta \leq 1$, then as $x \rightarrow \infty$,

$$\frac{1 - F(y)}{\gamma^\alpha c_\alpha (1 + \beta) x^{-\alpha}} \rightarrow 1, \frac{f(y)}{\alpha \gamma^\alpha c_\alpha (1 + \beta) x^{-(\alpha+1)}} \rightarrow 1 \quad (3.26)$$

$$\text{where } c_\alpha = \frac{\sin(\frac{\pi\alpha}{2})\Gamma(\alpha)}{\pi}$$

The lower tail properties are similar for $-1 \leq \beta < 1$ as $x \rightarrow \infty$:

$$\frac{F(-y)}{\gamma^\alpha c_\alpha (1 - \beta) x^{-\alpha}} \rightarrow 1, \frac{f(-y)}{\alpha \gamma^\alpha c_\alpha (1 - \beta) x^{-(\alpha+1)}} \rightarrow 1 \quad (3.27)$$

(ii) Stable distributions are unimodal as described previously.

(iii) Laws of stability have densities with uniformly bounded derivatives of every order.

Property 3.3.3 (Yang, 2012a)

Any admissible parameter quadruples $(\alpha, \beta_k, \gamma_k, \delta_k)$ and every real numbers h and $c_k, k = 1, \dots, n$, uniquely determine a parameter quadruple $(\alpha, \beta, \gamma, \delta)$ such that

$$S(\alpha, \beta, \gamma, \delta) \stackrel{d}{=} \sum_k c_k S(\alpha, \beta_k, \gamma_k, \delta_k) + h$$

With parameterization form A, the dependence of the quadruple $(\alpha, \beta, \gamma, \delta)$ on the chosen parameters and numbers is:

$$\delta = \sum_k \delta_k |c_k|^\alpha$$

$$\delta\beta = \sum_k \delta_k \beta_k |c_k|^\alpha \text{sign}(c_k)$$

$$\delta\gamma = \sum_k \delta_k \gamma_k c_k + h_0$$

where $h_0 = h$ if $\alpha \neq 1$ and $h_0 = h - \frac{2}{\pi} \sum_k \delta_k \beta_k c_k \log |c_k|$ if $\alpha = 1$.

Property 3.3.4 (Yang, 2012a)

Any two admissible parameter quadruples $(\alpha, \beta, \gamma, \delta)$ and $(\alpha, \beta, \gamma', \delta')$ uniquely determine real numbers $a > 0$ and b such that

$$S(\alpha, \beta, \gamma, \delta) \stackrel{d}{=} a S(\alpha, \beta, \gamma', \delta') + \lambda b,$$

With the parameterization in the A form, the dependence of a and b on the parameters is expressed as follows:

$$a = (\gamma/\gamma')^{1/\alpha} \tag{3.28}$$

$$b = \begin{cases} \delta - \delta' (\gamma/\gamma')^{\frac{1}{\alpha}-1}; & \alpha \neq 1, \\ \delta - \delta' + \frac{2}{\pi} \beta \log(\gamma/\gamma') & \alpha = 1. \end{cases} \tag{3.29}$$

This property is used to standardize any stable distribution by letting $\delta = 0$ and $\gamma = 1$.

3.4 Properties of stable laws

We summarize some basic properties of Nolan's S_1 -parameterization; that is $Y \sim S(\alpha, \beta, \gamma, \delta; 1)$ without proof.

- If $\beta = 0$, then it is implied that the stable distribution is symmetric.
- Reflection property is such that: $-Y \sim S(\alpha, -\beta, \gamma, -\delta; 1)$.
- All stable laws have densities $f(y)$ that are smooth and unimodal.
- The support of Y is the whole real line and exceptions exist when $\alpha < 1$ and $\beta = 1$, where the support is $[\delta, +\infty)$ or when $\alpha < 1$ and $\beta = -1$. In this case, the support is $(-\infty, \delta]$.
- Tail behavior: If $\alpha < 2$ and $-1 < \beta \leq 1$, then the density and distribution functions have an asymptotic power law. As $y \rightarrow \infty$,

$$1 - F(x) = P(X > x) \sim \gamma^\alpha c_\alpha (1 + \beta) x^{-\alpha} \quad (3.30)$$

$$f(x|\alpha, \beta, \gamma, \delta; 0) \sim \alpha \gamma^\alpha (1 + \beta) x^{-(\alpha+1)} \quad (3.31)$$

where $c_\alpha = \frac{\sin(\frac{\pi\alpha}{2})\Gamma(\alpha)}{\pi}$. *Stable Paretian distribution* is used in the non-Gaussian case owing to the similarity of tail behavior with the Pareto distribution $\forall \alpha < 2$ and $-1 < \beta < 1$, both tail probabilities and densities are asymptotically power laws. In the case when $\beta = -1$, the right tail of the distribution is not asymptotically power law. In the same way, when $\beta = 1$ the left tail is not asymptotically power law.

- The Generalized Central Limit Theorem is also a basic property of stable laws and is discussed in detail in Appendix A.
- Fractional moments: When $\alpha < 2$, $E|X|^p$ is finite for $0 < p < \alpha$, but infinite for $p \geq \alpha$. For $\alpha < 2$, the population variance is infinite and for $\alpha \leq 1$, the population mean is undefined. This is a consequence of the power law tail behavior (Nolan, 2014).

3.5 Sum of stable random variables

The basic property of stable laws is that sums of α –stable random variables are α –stable. However, results depend on the parameterization used.

Property 3.5.1 (Nolan, 2015)

The $S(\alpha, \beta, \gamma, \delta; 0)$ have the following properties:

- (a) If $Y \sim S(\alpha, \beta, \gamma, \delta; 0)$, then for any $a \neq 0$ and $b \in \mathbb{R}$,

$$aY + b \sim S(\alpha, (\text{sign } a)\beta, |a|\gamma, a\delta + b; 0)$$

- (b) The characteristic density, distribution and characteristic functions are jointly continuous in all four parameters $(\alpha, \beta, \gamma, \delta)$.

- (c) If $Y_1 \sim S(\alpha, \beta_1, \gamma_1, \delta_1; 0)$ and $Y_2 \sim S(\alpha, \beta_2, \gamma_2, \delta_2; 0)$ are independent, then $Y_1 + Y_2 \sim S(\alpha, \beta, \gamma, \delta; 0)$ where

$$\beta = \frac{\beta_1 \gamma_1^\alpha + \beta_2 \gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}, \gamma_\alpha = \gamma_1^\alpha + \gamma_2^\alpha$$

$$\delta = \begin{cases} \delta_1 + \delta_2 + \tan \frac{\pi\alpha}{2} [\beta\gamma - \beta_1\gamma_1 - \beta_2\gamma_2]; & \alpha \neq 1, \\ \delta_1 + \delta_2 + \frac{2}{\pi} [\beta\gamma \log \gamma - \beta_1\gamma_1 \log \gamma_1 - \beta_2\gamma_2 \log \gamma_2] & \alpha = 1. \end{cases}$$

$\gamma_\alpha = \gamma_1^\alpha + \gamma_2^\alpha$ is the general rule for addition of variances of independent random variables and holds for both parameterizations.

Property 3.5.2 (Nolan, 2015)

The $S(\alpha, \beta, \gamma, \delta; 1)$ have the following properties:

- (a) If $Y \sim S(\alpha, \beta, \gamma, \delta; 1)$ then for any $a \neq 0$ and $b \in \mathbb{R}$,

$$aY + b \sim \begin{cases} S(\alpha, (\text{sign } a)\beta, |a|\gamma, a\delta + b; 1) & \alpha \neq 1, \\ S(1, (\text{sign } a)\beta, |a|\gamma, a\delta + b - \frac{2}{\pi} \beta \gamma a \log(|a|); 1) & \alpha = 1. \end{cases}$$

- (b) The characteristic density, distribution and characteristic functions are continuous away from $\alpha = 1$ and discontinuous in any neighborhood of $\alpha = 1$.

(c) If $Y_1 \sim S(\alpha, \beta_1, \gamma_1, \delta_1; 1)$ and $Y_2 \sim S(\alpha, \beta_2, \gamma_2, \delta_2; 1)$ are independent, then $Y_1 + Y_2 \sim S(\alpha, \beta, \gamma, \delta; 1)$ where

$$\beta = \frac{\beta_1 \gamma_1^\alpha + \beta_2 \gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}, \quad \gamma_\alpha = \gamma_1^\alpha + \gamma_2^\alpha, \quad \delta = \delta_1 + \delta_2$$

(Nolan, 2015).

Property 3.5.1 shows that γ and δ are the standard scale and location parameters in the Nolan's S_0 -parameterization but not in the Nolan's S_1 -parameterization in the case when $\alpha = 1$. Part (c) in the first parameterization shows that δ (the location parameter) is the sum $\delta_1 + \delta_2$.

By induction, we may generate formulas for the sum of n stable random variables.

For $Y_j \sim S(\alpha, \beta_j, \gamma_j, \delta_j; k), j = 1, 2, \dots, n$, and independent and arbitrary w_1, \dots, w_n , the sum

$$w_1 Y_1 + w_2 Y_2 + \dots + w_n Y_n \sim S(\alpha, \beta, \gamma, \delta; k) \quad (3.32)$$

where

$$\gamma^\alpha = \sum_{j=1}^n |w_j \gamma_j|^\alpha$$

$$\beta = \frac{\sum_{j=1}^n \beta_j (\text{sign}(w_j)) |w_j \gamma_j|^\alpha}{\gamma^\alpha}$$

$$\delta = \begin{cases} \sum_j w_j \delta_j + \tan \frac{\pi \alpha}{2} (\beta \gamma - \sum_j \beta_j w_j \gamma_j) & k = 0, \alpha \neq 1, \\ \sum_j w_j \delta_j + \frac{2}{\pi} (\beta \gamma \log \gamma - \sum_j \beta_j w_j \gamma_j \log |w_j \gamma_j|) & k = 0, \alpha = 1, \\ \sum_j w_j \delta_j & k = 1, \alpha \neq 1, \\ \sum_j w_j \delta_j - \frac{2}{\pi} \sum_j \beta_j w_j \gamma_j \log |w_j| & k = 1, \alpha = 1. \end{cases}$$

If $\beta_j = 0 \forall j$, then $\beta = 0$ and $\delta = \sum_j w_j \delta_j$. We further note an important property called the scaling property for random variables.

When terms are independent and identically distributed,

$$Y_j \sim S(\alpha, \beta, \gamma, \delta; k),$$

then,

$$Y_1 + \dots + Y_n \sim S(\alpha, \beta, n^{1/\alpha}\gamma, \delta_n; k) \quad (3.33)$$

where

$$\delta_n = \begin{cases} n\delta + \gamma\beta \tan \frac{\pi\alpha}{2} (n^{1/\alpha} - n) & k = 0, \alpha \neq 1, \\ n\delta + \gamma\beta \frac{\pi\alpha}{2} n \log n & k = 0, \alpha = 1, \\ n\delta & k = 1. \end{cases}$$

The shape of the sum of n terms remains the same as the original shape. It is to be pointed out that no other distribution has this property of stable distributions (Nolan, 2015).

3.6 Stable parameter estimation

3.6.1 Univariate estimation

Nolan (2015) states that many standard parameter estimation procedures fail to work for stable data since there is a lack of closed-form densities for stable distributions as discussed earlier. The very common method of moments where one is first required to compute $E[X]$, $E[X^2]$, $E[X^3]$ and $E[X^4]$ to estimate the four stable parameters, that is solve for α, β, γ and δ is not applicable as all of these moments do not exist. However, sample moments do exist but their behavior is erratic. The likelihood function cannot be expressed explicitly by the argument of the lack of closed analytic form stable densities. This causes difficulties in solving for maximum likelihood estimators. There are, however, many non-standard procedures for estimating stable parameters. We describe these methods below:

- Fama & Roll (1968) proposed the oldest method of estimating stable parameters known as *Quantile Matching*. They noticed certain patterns in quantiles x_p (p -th quantile of a distribution) of a symmetrically stable distribution ($\beta = 0$) which could be used to estimate α and the γ . These ideas were further developed by McCulloch (1986) for the general asymmetric case where bias is removed and consistent estimates for all four parameters are obtained. This method uses five sample quantiles, the 5^{th} , 25^{th} , 50^{th} , 75^{th} and 95^{th} quantiles. These values are matched to a stable distribution with the closest spread pattern. Reliable estimates may be obtained by this method if the sample set is large and the data set stable.
- Koutrouvelis (1980) used the *empirical characteristic function method*, where there is an explicit formula, equation (3.13) for the characteristic function $\phi(t)$. The sample or empirical characteristic function $\hat{\phi}(u_i)$ can be computed on a grid of u_i values for a given data set and then uses regression analysis to estimate the parameters. A simplified method was identified by Kogon & Williams (1998). This method used the continuous parameterization, equation (3.11) and then centering and scaling the data to avoid possible numerical difficulties.
- Buckle (1995) proposed a *Bayesian inference* method for estimation by using MCMC method with the Gibbs sampler. The posterior density in Bayesian inference can be

found from:

$$\pi(\theta|y) \propto f(y|\theta)\pi(\theta).$$

That is, posterior \propto likelihood \times prior. In the case where the probability density function of y is unobtainable in closed-form whereas the joint pdf of y and z exist, then the posterior density is found by using the integration,

$$\pi(\theta|y) \propto \int f(y, z|\theta)\pi(\theta)dz$$

(Oral et al., 2012).

- Nikias & Shao (1995) used a method of estimation known as the *fractional and negative method of moments*. This method is used for symmetric stable distributions where $\beta = 0$ and $\delta = 0$. When Y is strictly stable, there exists expressions for fractional moments $E|Y|^p$, for $-1 < p < \alpha$. One can use the above expression for a generalized method of moments, where one is required to compute sample fractional moments, set them equal to the expressions in terms of the parameters and solve for each of the parameters.
- *Tail estimation* is a method that uses the tail behavior

$$1 - F(x) = P(X > x) \sim \gamma^\alpha c_\alpha (1 + \beta)x^{-\alpha} \quad (3.34)$$

$$f(x|\alpha, \beta, \gamma, \delta; 0) \sim \alpha\gamma^\alpha (1 + \beta)x^{-(\alpha+1)} \quad (3.35)$$

where $c_\alpha = \frac{\sin(\frac{\pi\alpha}{2})\Gamma(\alpha)}{\pi}$ to estimate α . Different methods have been suggested. The Hill estimator and generalizations to plotting extremes on a log-log scale followed by estimating the slope. Albeit, these do not work well with stable laws as when the power law occurs, it is a complicated function of the parameters. Unless one has a fairly large data set, it is highly unlikely that the tail will be exactly a power law (Nolan, 2015).

- The *maximum likelihood (ML) estimation method* is the most commonly used method in stable parameter estimation. This method is discussed in more detail in the next section.

A simulation study by Ojeda (2001) found that the ML method yields the most accurate

results followed by the empirical characteristic function method, thereafter by the quantile method, and lastly, the fractional moment method (Nolan, 2003).

3.6.2 Maximum likelihood estimation

The parameter vector is denoted by $\vec{\theta} = (\alpha, \beta, \gamma, \delta_0)$ and the density function is denoted by $f(x|\vec{\theta})$. $\Theta = (0, 2] \times [-1, 1] \times (0, \infty) \times (-\infty, \infty)$ denotes the parameter space. The log-likelihood function for an independent and identically distributed stable sample Y_1, \dots, Y_n is given by

$$L(\vec{\theta}) = \sum_{i=1}^n \log f(Y_i|\vec{\theta})$$

Since there are no closed formulas for general stable densities, there are some difficulties trying to compute the likelihood function. The program STABLE computes stable densities that are reliable for $\alpha > 0.1$ and any β, γ and δ_0 . The McCulloch (1986) quantile method is used initially to approximate the parameters and the parameter space can constrain a method to maximize called the quasi-Newton method. DuMouchel (1971, 1973) indicated that if $\vec{\theta}_0$ lies in the interior of the parameter space Θ , the maximum likelihood estimator is consistent and asymptotically Normal with mean $\vec{\theta}_0$ and covariance matrix given by $n^{-1}B$ where $B = (b_{ij})$ is the inverse of the 4×4 Fisher information matrix \mathbf{I} . Entries in \mathbf{I} are given by

$$\mathbf{I}_{ij} = \int_{-\infty}^{\infty} \frac{\partial f}{\partial \theta_i} \frac{\partial f}{\partial \theta_j} dy$$

The behavior of the estimators is unknown when $\vec{\theta}$ is near the boundary of the parameter space. The distribution of the estimator gets skewed away from the boundary. When $\alpha = 2$ or $\beta = \pm 1$, $\vec{\theta}$ is on the boundary of the parameter space. The Normal distribution for the estimators tends to a degenerate distribution at the boundary point. Away from the boundary, large sample confidence intervals for each of the parameters are given by

$$\hat{\theta}_i \pm Z_{\frac{\alpha}{2}} \frac{\sigma_{\hat{\theta}_i}}{\sqrt{n}}$$

Where $\sigma_{\hat{\theta}_1} \cdots \sigma_{\hat{\theta}_n}$ are the diagonal entries of B (Nolan, 2001).

Chapter 4

Multivariate stable distributions

In this chapter, we explore multivariate stable laws with specific interest in elliptically contoured stable distributions.

4.1 Multivariate stable distributions

Stable laws can be extended to multidimensional cases. In this section, we consider a sequence of independent and identically distributed random variables Y_1, Y_2, Y_3, \dots in the n -dimensional Euclidean space \mathbb{R}^n and form the sequence of sums

$$Z_n = \frac{(Y_1 + \dots + Y_n - a_n)}{b_n}, \quad n = 1, 2, \dots$$

normalized by some sequences of positive numbers b_n and non-random elements $a_n \in \mathbb{R}^n$. Alternatively, if the sums $Y_1 + Y_2 + \dots + Y_n$ are normalized by non-singular matrices Σ_n and not by positive numbers $\frac{1}{b_n}$, then concepts of stable distributions become more versatile. At present, there is limited knowledge about the analytical properties of multivariate stable laws which contrasts greatly from the vast amount of facts known from univariate stable distributions. We look at the canonical representation of the characteristic function $t_N(k), k \in \mathbb{R}^n$, of finite dimensional Lévy-Feldheim laws.

The characteristic functions are of the form

$$t_N = e^{i(k,a) - \psi_\alpha(k)}, \quad 0 < \alpha \leq 2 \quad (4.1)$$

where $a \in \mathbb{R}^n$ and the functions $\psi_\alpha(k)$ which are determined by the parameter α and by the finite measure $M(du)$ on the sphere $\mathbb{S} = \{u : |u| = 1\}$. If $\alpha = 2$, then $\psi_\alpha(k) = (\Sigma k, k)$ where Σ is the so-called covariance matrix.

If $0 < \alpha < 2$, then

$$\psi_\alpha(k) = \int_S |(k, u)|^\alpha \omega_\alpha(k, u) M(du), \quad (4.2)$$

where

$$\omega_\alpha(k, u) = \begin{cases} 1 - i \tan \frac{\pi\alpha}{2} \text{sign}(k, u); & \alpha \neq 1, \\ 1 + i \left(\frac{2}{\pi}\right) \ln |(k, u)| \text{sign}(k, u) & \alpha = 1. \end{cases} \quad (4.3)$$

$$\ln t_N(k) = \begin{cases} \lambda [i |k|^\gamma - |k|^\alpha (1 - i \beta \tan(\frac{\pi\alpha}{2}))]; & \alpha \neq 1, \\ \lambda [i |k|^\gamma - |k| (1 - i \frac{2}{\pi} \beta \ln |k|)]; & \alpha = 1. \end{cases} \quad (4.4)$$

where $0 < \alpha \leq 2$ and β, γ and λ are real-valued functions defined on the unit sphere \mathbb{S} determined by:

$$\lambda = \lambda(u) = \int_S |(u, u')|^\alpha M(du'), \quad u \in S$$

$$\lambda\beta = \lambda\beta(u) = \int_S |(u, u')|^\alpha \text{sign}(u, u') M(du'),$$

$$\lambda\gamma = \lambda\gamma(u) = \begin{cases} (u, a), & \alpha \neq 1, \\ (u, a) - \left(\frac{2}{\pi}\right) \int_S (u, u') \ln |(u, u')| M(du') & \alpha = 1, \end{cases}$$

We give some properties of the functions β, γ and λ .

- (a) β, γ and λ are continuous on \mathbb{S} and for a given α , they determine a unique shift in a and the measure $M(du)$ in equations (4.1) and (4.2). When $\alpha \neq 1$, the functions β and λ determine the measure M .
- (b) The domain is given by the function γ , that is, the entire real axis.
- (c) The following holds for any $u \in S$:

$$\beta(-\mu) = -\beta(u), \quad \lambda(-u) = \lambda(u),$$

$$|\beta(u)| \leq 1, \quad 0 \leq \lambda(u) \leq M_0,$$

where M_0 is the value of the complete measure M on \mathbb{S} .

All inequalities are strict unless $M(du)$ is concentrated entirely on some subspace of \mathbb{R}^n . Then,

$$\lambda_0 = \inf \lambda(u) : u \in \mathbb{S} > 0, \quad |t_N(k)| \leq e^{(\lambda_0|k|^\alpha)}$$

Hence, the corresponding stable distribution denoted by (Uchaikin & Zolotarev, 1999) has density $q_N(y; \alpha, a, M)$ bounded by

$$\frac{\Gamma(1 + \frac{N}{\alpha})}{\Gamma(1 + \frac{N}{2})} (2\sqrt{\pi}\lambda^{\frac{1}{\alpha}})^{-n}$$

4.1.1 Multivariate stable laws

Feldheim (1937) showed that every multivariate stable vector has a characteristic function:

$$\phi(u) = E[e^{i\mathbf{u} \cdot Y}] = \exp \left(\int_{\mathbb{S}} \omega_\alpha(\mathbf{u} \cdot \mathbf{s}) \Lambda(ds) + i\mathbf{u} \cdot \delta \right)$$

where Λ is a finite measure on a unit sphere $\mathbb{S} = \{|\mathbf{y}| = 1\}$, δ is a shift vector in \mathbb{R}^d and

$$\omega_\alpha(t) = -\log E[e^{itZ}] = \begin{cases} |t|^\alpha [1 - \tan \frac{\pi\alpha}{2} (\text{sign})(t)]; & \alpha \neq 1, \\ |t| [1 + i \frac{2}{\pi} (\text{sign}(t)) \log |t|] & \alpha = 1. \end{cases}$$

$\omega_\alpha(t)$ is a subtraction of the exponent of the characteristic function of a univariate $Z \sim S(\alpha, \beta = 1, \gamma = 1, \delta = 0; 1)$. Every multivariate stable law is characterized by α and a spectral measure Λ on the sphere and a shift vector δ (Nolan, 2014).

A d -dimensional α -stable random vector is determined by a spectral measure Γ - a finite Borel measure on $\mathbb{S}_d = \{\mathbf{s} : \|\mathbf{s}\| = 1\}$ is the unit sphere in \mathbb{R}^d . The notation $\mathbf{Y} \sim S_{\alpha,d}(\Gamma, \mu^0)$ is used to denote a stable random vector in this case.

The characteristic function of $\mathbf{Y} \sim S_{\alpha,d}(\Gamma, \mu^0)$ is described by:

$$\phi_{\mathbf{Y}}(t) = E[e^{i\langle \mathbf{Y}, \mathbf{t} \rangle}] = e^{(-I_{\mathbf{Y}}(\mathbf{t}) + i\langle \mu^0, \mathbf{t} \rangle)} \quad (4.5)$$

where the function $I_{\mathbf{Y}}(\mathbf{t})$ in the exponent is

$$I_{\mathbf{Y}}(\mathbf{t}) = \int_{\mathbb{S}_d} \psi_{\alpha}(\langle \mathbf{t}, \mathbf{s} \rangle) \Gamma(ds). \quad (4.6)$$

Here $\langle \mathbf{t}, \mathbf{s} \rangle = t_1 s_1 + \dots + t_d s_d$ is the inner product and

$$\psi_{\alpha}(u) = \begin{cases} |u|^{\alpha} (1 - i \tan \frac{\pi \alpha}{2} \text{sign}(u)); & \alpha \neq 1, \\ |u| (1 + i (\frac{2}{\pi}) \ln |u| \text{sign}(u)) & \alpha = 1. \end{cases}$$

The characteristic function is determined by $I_{\mathbf{Y}}(\mathbf{t})$, $\mathbf{t} \in \mathbb{R}^d$ and the complex valued function $I_{\mathbf{Y}}(t)$ determines the distribution of \mathbf{Y} . Below, we use this concept to estimate Γ from the data.

For any $\mathbf{t} \in \mathbb{R}^d$, the projection of the random vector $\langle \mathbf{t}, \mathbf{X} \rangle$ is a one-dimensional random variable with a characteristic function given by: $E[e^{iu\langle \mathbf{t}, \mathbf{Y} \rangle}] = e^{-I_{\mathbf{Y}}(u\mathbf{t})}$. Therefore, the scale, skewness and shift functions are given by Zolotarev (1986) :

$$\begin{aligned} \sigma^{\alpha}(\mathbf{t}) &= \Re I_{\mathbf{X}}(\mathbf{t}) = \int_{\mathbb{S}_d} |\langle \mathbf{t}, \mathbf{s} \rangle|^{\alpha} \Gamma(ds) \\ \beta(\mathbf{t}) &= \sigma^{-\alpha}(\mathbf{t}) \int_{\mathbb{S}_d} \text{sign} |\langle \mathbf{t}, \mathbf{s} \rangle|^{\alpha} \Gamma(ds) = \begin{cases} \Im \frac{I_{\mathbf{Y}}(\mathbf{t})}{\sigma^{\alpha}(\mathbf{t}) \tan \frac{\pi \alpha}{2}} & \alpha \neq 1, \\ \Im \left[\frac{I_{\mathbf{Y}}(2\mathbf{t}) - 2I_{\mathbf{Y}}(\mathbf{t})}{4\sigma(\mathbf{t}) \ln \frac{2}{\pi}} \right] & \alpha = 1 \end{cases} \\ \mu(\mathbf{t}) &= \begin{cases} \langle \mathbf{t}, \mu^0 \rangle & \alpha \neq 1, \\ \langle \mathbf{t}, \mu^0 \rangle - \frac{2}{\pi} \int \langle \mathbf{t}, \mathbf{s} \rangle \ln |\langle \mathbf{t}, \mathbf{s} \rangle| \Gamma(ds) = \Im \frac{I_{\mathbf{Y}}(\mathbf{t})}{\sigma(\mathbf{t})} & \alpha = 1. \end{cases} \end{aligned} \quad (4.7)$$

We take $\mu^0 = 0$ by replacing \mathbf{Y} with $\mathbf{Y} - \mu^0$. The functions $\sigma(\mathbf{t})$, $\beta(\mathbf{t})$ and $\mu(\mathbf{t})$ determine the distribution. A relatively easy way to see this is to determine $I_{\mathbf{Y}}(\cdot)$:

$$I_{\mathbf{Y}} = \begin{cases} \sigma^\alpha(\mathbf{t})(1 - i\beta(\mathbf{t})\tan \frac{\pi\alpha}{2}) & \alpha \neq 1, \\ \sigma(\mathbf{t})(1 - i\mu(\mathbf{t})) & \alpha = 1. \end{cases}$$

Another way of describing a stable random vector is in terms of projections. For any vector \mathbf{u} , the projection $\langle \mathbf{u}, \mathbf{Y} \rangle$ is univariate α -stable with skewness $\beta(\mathbf{u})$, scale $\gamma(\mathbf{u})$ and shift $\delta(\mathbf{u})$. We write $\mathbf{Y} \sim S(\alpha, \beta(\cdot), \gamma(\cdot), \delta(\cdot))$ if \mathbf{Y} is stable with

$$\langle \mathbf{u}, \mathbf{Y} \rangle \sim S(\alpha, \beta(\mathbf{u}), \gamma(\mathbf{u}), \delta(\mathbf{u}))$$

for every $\mathbf{u} \in \mathbb{R}^d$. This is known as the *projection parameterization*.

The spectral measure determines projection parameter functions described below by:

$$\begin{aligned} \gamma(\mathbf{u}) &= \left(\int_{\mathbb{S}} |\langle \mathbf{u}, \mathbf{s} \rangle|^\alpha \Lambda(ds) \right)^{1/\alpha} \\ \beta(\mathbf{u}) &= \frac{\int_{\mathbb{S}} |\langle \mathbf{u}, \mathbf{s} \rangle|^\alpha \text{sign}(\langle \mathbf{u}, \mathbf{s} \rangle) \Lambda(ds)}{\gamma(\mathbf{u})} \\ \delta(\mathbf{u}) &= \begin{cases} \langle \mathbf{u}, \delta \rangle & \alpha \neq 1 \\ \langle \mathbf{u}, \delta \rangle - \frac{2}{\pi} \int_{\mathbb{S}} \langle \mathbf{u}, \mathbf{s} \rangle \ln |\langle \mathbf{u}, \mathbf{s} \rangle| \Lambda(ds) & \alpha = 1. \end{cases} \end{aligned}$$

4.2 Discrete spectral measures

This section discusses the case when Γ is a discrete spectral measure described by a finite number of point masses:

$$\Gamma(\cdot) = \sum_{j=1}^n \gamma_j \delta_{\mathbf{s}_j}(\cdot) \quad (4.8)$$

where γ_j 's are the weights and $\delta_{\mathbf{s}_j}$'s are point masses at the points $\mathbf{s}_j \in \mathbb{S}_d, j = 1, \dots, n$. When the components \mathbf{Y} are independent or when \mathbf{Y} arises from the finite dimensional distributions of a stable Ornstein-Uhlenbeck process and when one estimates Γ from the data, then such spectral measures arise naturally in these several cases. Discrete spectral measures are a simple class to study. We explain what is meant by “dense” spectral measures.

For a discrete spectral measure given by equation (4.8), the characteristic equation (4.5) becomes

$$\phi^*(\mathbf{t}) = \exp \left(- \sum_{j=1}^n \psi_\alpha(\langle \mathbf{t}, \mathbf{s}_j \rangle) \gamma_j \right). \quad (4.9)$$

The above expression is numerically simple to compute whereas $\phi(\mathbf{t})$ is more difficult to compute. Let p be the density function corresponding to equation (4.5) and let p^* be the density with characteristic function of equation (4.8).

Theorem 4.2.1 (Byczkowski et al., 1993)

Let \mathbf{Y} be a truly d -dimensional α -stable random vector ($d \geq 2, 0 < \alpha < 2$) with zero shift, spectral measure σ and density $p(\mathbf{y})$. Let $\epsilon > 0$, then

- (i) There exists a discrete measure σ^* with corresponding stable density $p^*(\mathbf{y})$ satisfying

$$\sup_{\mathbf{y} \in \mathbb{R}^d} |p(\mathbf{y}) - p^*(\mathbf{y})| \leq \epsilon.$$

- (ii) There is a discrete measure σ' with corresponding stable random vector \mathbf{Y}' which satisfies

$$\sup_{A \in \text{Borel}(\mathbb{R}^d)} |P(\mathbf{Y} \in A) - P(\mathbf{Y}' \in A)| \leq \epsilon$$

Remark. (Byczkowski et al., 1993)

We define the discrete measure by considering:

A finite partition A_1, \dots, A_n of \mathbb{S}^{d-1} , points $\mathbf{s}_1, \dots, \mathbf{s}_n \in \mathbb{S}^{d-1}$ and σ^* is defined by concentrating mass $\sigma(A_j)$ at \mathbf{s}_j . that is,

$$\sigma^*(\cdot) = \sum_{j=1}^n \sigma(A_j) \delta_{\mathbf{s}_j}(\cdot)$$

Lemma 4.2.1 (Byczkowski et al., 1993)

Using notation from Theorem 4.2.1, for any $R > 0, \epsilon > 0$, there is a discrete measure σ^* on \mathbb{S}^{d-1} with at most $2 \left[\pi \sqrt{d-1} / 2\delta(\epsilon/\|\sigma\| : \alpha, R) \right]^{d-1}$ atoms which satisfies $\sigma^{**}(\mathbb{S}^{d-1}) = \|\sigma\|$ and

$$\sup_{|\mathbf{t}| \leq R} |I(\mathbf{t}) - I^*(\mathbf{t})| \leq \epsilon, \quad (4.10)$$

where

$$\delta(\epsilon; \alpha, R) = \begin{cases} (1 + 2 \left| \tan \frac{\pi\alpha}{2} \right|)^{-1/\alpha} \epsilon^{1/\alpha} & 0 < \alpha \leq 1, \\ \min(e^{\alpha-1}, \epsilon/2(1 + \ln_+ R), \pi\epsilon/16\ln_+(\frac{1}{\epsilon})) & \alpha = 1, \\ \alpha^{-1} R^{1-\alpha} (1 + \left| \tan \frac{\pi\alpha}{2} \right|)^{-1} \epsilon & 1 < \alpha < 2. \end{cases}$$

Lemma[4.2.2] (Byczkowski et al., 1993)

Let $\mathbf{Y} = (Y_1, \dots, Y_d)$ be a d -dimensional α -stable random vector with spectral measure σ and zero shift parameter. For $\epsilon > 0, P(|\mathbf{Y}| > R \leq \epsilon)$ whenever R satisfies

$$R^{\alpha/2} \geq \begin{cases} c(\frac{\alpha}{2}, \alpha, 1) d \|\sigma\|^{1/\alpha} / \epsilon & 0 < \alpha \neq 1 \text{ or } (\alpha = 1 \text{ and } \sigma \text{ symmetric}) \\ (\frac{1}{2}, \alpha, 1) d(2 + \|\sigma\| + \sqrt{2\|\sigma\|/\pi\epsilon}) / \epsilon & \alpha = 1 \text{ and } \sigma \text{ nonsymmetric} \end{cases} \quad (4.11)$$

where for any $0 < p < \alpha$ and skewness parameter β

$$c(p, \alpha, \beta) = \begin{cases} \frac{(2^{p-1} \Gamma(1 - (\frac{p}{\alpha})) (1 + \beta^2 \tan^2(\frac{\pi\alpha}{2}))^{p/2\alpha} \times \cos((\frac{\pi}{\alpha}) \arctan(\beta \tan(\frac{\pi\alpha}{2})))}{p \int_0^\infty u^{p-1} \sin^2 u du} & \alpha \neq 1 \\ & \text{or } (\alpha = 1, \beta = 0) \\ 2\Gamma(p) \frac{\sin(\frac{\pi\alpha}{2})}{\pi(1-p)} & \alpha = 1 \\ & \text{and } \beta \neq 0 \end{cases}$$

4.3 Multivariate stable parameter estimation

4.3.1 Estimating discrete spectral measures

In stochastic modeling of financial portfolios, the need for an estimator of a multivariate α -stable spectral measure arises. The spectral measure carries essential information about the vector. There are three solutions to the estimation problem namely the Rachev-Xin-Cheng based on Theorem 4.2.1, the empirical characteristic function (ECF) method and the method devised by McCulloch (1994) called the projection method.

Rachev-Xin-Cheng method Rachev & Xin (1993) and Cheng & Rachev (1995):

An *ad hoc* is chosen for r which is used to estimate the measure of set $A \subset \mathbb{S}_d$ by:

$$\hat{\Gamma}(A) = \text{const.} \frac{\#\{\mathbf{Y}_i : |\mathbf{Y}_i| > r, \mathbf{Y}_i \in \text{Cone}(\mathbb{A})\}}{\#\{\mathbf{Y}_i : |\mathbf{Y}_i| > r\}}$$

The next two estimators are based on using a sample to estimate the characteristic function on some grid. In particular, we estimate the exponent of the characteristic function $I_{\mathbf{Y}}(\cdot)$ on a grid $\mathbf{t}_1, \dots, \mathbf{t}_n \in \mathbb{S}^d$.

Empirical characteristic function method (ECF)

This method is fairly straightforward. Consider an i.i.d. sample $\mathbf{Y}_1, \dots, \mathbf{Y}_k$ of α -stable random vectors with the spectral measure Γ . Let $\hat{\phi}_k(\mathbf{t})$ and \hat{I}_k be the empirical counterparts of ϕ and I ; that is $\hat{\phi}_k(\mathbf{t}) = (\frac{1}{k}) \sum_{i=1}^k e^{i\langle \mathbf{t}, \mathbf{Y}_i \rangle}$ is the sample characteristic function and $\hat{I}_k(\mathbf{t}) = -\ln \hat{\phi}_k(\mathbf{t})$. Given a grid $\mathbf{t}_1, \dots, \mathbf{t}_n \in \mathbb{S}^d$, then $\vec{I}_{ECF,k} = [\hat{I}_k(\mathbf{t}_1), \dots, \hat{I}_k(\mathbf{t}_n)]'$ is the ECF estimate of $I_{\mathbf{Y}}(\cdot)$. The average coordinate of α 's, $\alpha_{ECF} = \bar{\alpha}$, where $\bar{\alpha} = \frac{(\sum_{j=1}^d \hat{\alpha}_j)}{d}$ is defined as the estimate of the joint index of stability α .

Projection method

This method is based on one-dimensional projections of the data set. Consider a projection of the random vector \mathbf{Y} for any $\mathbf{t} \in \mathbb{S}^d$, $\langle \mathbf{t}, \mathbf{Y} \rangle$ is a one-dimensional random variable with the characteristic function:

$$E[e^{iu\langle \mathbf{t}, \mathbf{Y} \rangle}] = e^{-I_{\mathbf{Y}}(u\mathbf{t})}.$$

Zolotarev (1986) and Samoradnitsky & Taquq (1994) describe its scale, skewness and shift by

$$\begin{aligned}\sigma^\alpha(\mathbf{t}) &= \Re I_{\mathbf{y}}(\mathbf{t}) = \int_{\mathbb{S}^d} |\langle \mathbf{t}, \mathbf{s} \rangle|^\alpha \Gamma(d\mathbf{s}) \\ \beta(\mathbf{t}) &= \sigma^{-\alpha}(\mathbf{t}) \int_{\mathbb{S}^d} \text{sign} \langle \mathbf{t}, \mathbf{s} \rangle |\langle \mathbf{t}, \mathbf{s} \rangle|^\alpha \Gamma(d\mathbf{s}) \\ &= \begin{cases} \frac{-\Im I_{\mathbf{y}}(\mathbf{t})}{\sigma^\alpha(\mathbf{t}) \tan(\frac{\pi\alpha}{2})} & \alpha \neq 1, \\ \frac{\Im[I_{\mathbf{y}}(2\mathbf{t}) - 2I_{\mathbf{y}}(\mathbf{t})]}{4\sigma(\mathbf{t}) \ln(\frac{2}{\pi})} & \alpha = 1. \end{cases} \\ \mu(\mathbf{t}) &= \begin{cases} 0 & \alpha \neq 1, \\ -\frac{2}{\pi} \int \langle \mathbf{t}, \mathbf{s} \rangle \ln |\langle \mathbf{t}, \mathbf{s} \rangle| \Gamma(d\mathbf{s}) = \frac{-\Im I_{\mathbf{y}}}{\sigma(\mathbf{t})} & \alpha = 1. \end{cases}\end{aligned}$$

Consider the sample $\mathbf{Y}_1, \dots, \mathbf{Y}_k$. Now fix a grid $\mathbf{t}_1, \dots, \mathbf{t}_n$ on \mathbb{S}^d , and for each \mathbf{t}_j define the one-dimensional data set $\langle \mathbf{t}_j, \mathbf{Y}_1 \rangle, \dots, \langle \mathbf{t}_j, \mathbf{Y}_k \rangle$. Use this method to estimate the scale $\hat{\sigma}(\mathbf{t}_j)$ and skewness $\hat{\beta}(\mathbf{t}_j)$ as well as the shift $\hat{\mu}(\mathbf{t}_j)$ when $\alpha = 1$ of this one-dimensional data. We define

$$\hat{I}_k(\mathbf{t}_j) = \begin{cases} \sigma^\alpha(\mathbf{t}_j) 1 - i\hat{\beta}(\mathbf{t}_j) \tan(\frac{\pi\alpha}{2}) & \alpha \neq 1, \\ \hat{\sigma}(\mathbf{t}_j) (1 - i\hat{\mu}(\mathbf{t}_j)) & \alpha = 1. \end{cases}$$

The vector, $\vec{I}_{PROJ,k} = [\hat{I}_k(\mathbf{t}_1), \dots, \hat{I}_k(\mathbf{t}_n)]'$ is the projection estimator of $I_{\mathbf{X}}(\cdot)$. In view of the fact that we estimated the parameters of each projection, we get an estimate $\hat{\alpha}(\mathbf{t}_j)$ for each direction. Also, for this method the average $\alpha_{PROJ} = \frac{1}{n} \sum \hat{\alpha}(\mathbf{t}_j)$ as a pooled estimate of α (Nolan et al., 2001).

In order to obtain the estimate of the spectral measure $\hat{\Gamma}$, there is a need to invert the discrete approximations to the characteristic function obtained by the ECF and PROJ method. Starting with the case when Γ is a discrete spectral measure of the form in equation (4.8). We let $I_{\mathbf{X}}(\mathbf{t}) = \sum_{j=1}^n \psi_\alpha(\langle \mathbf{t}, \mathbf{s}_j \rangle) \gamma_j$. Furthermore, let $\mathbf{t}_1, \dots, \mathbf{t}_n \in \mathbb{R}^d$. Define the $n \times n$ matrix

$$\Psi = \Psi(\mathbf{t}_1, \dots, \mathbf{t}_n; \mathbf{s}_1, \dots, \mathbf{s}_n) = \begin{bmatrix} \psi_\alpha(\langle \mathbf{t}_1, \mathbf{s}_1 \rangle) & \dots & \psi_\alpha(\langle \mathbf{t}_1, \mathbf{s}_n \rangle) \\ \vdots & \dots & \vdots \\ \psi_\alpha(\langle \mathbf{t}_n, \mathbf{s}_1 \rangle) & \dots & \psi_\alpha(\langle \mathbf{t}_n, \mathbf{s}_n \rangle) \end{bmatrix}$$

If $\vec{\gamma} = [\gamma_1, \dots, \gamma_n]'$ and $\vec{I} = [I_{\mathbf{X}}(\mathbf{t}_1), \dots, I_{\mathbf{X}}(\mathbf{t}_n)]'$, then

$$\vec{\mathbf{I}} = \vec{\Psi} \vec{\gamma}.$$

If $\mathbf{t}_1, \dots, \mathbf{t}_n \in \mathbb{R}^d$ are chosen so that Ψ^{-1} exists, then $\vec{\gamma} = \Psi^{-1} \vec{\mathbf{I}}$ is the exact solution to $\vec{\mathbf{I}}$ above.

For, Γ , the general spectral measure (which is not discrete and the location of the point masses are not known), we consider a discrete approximation $\Gamma^* = \sum_{j=1}^n \gamma_j \delta_{\mathbf{s}_j}$, where $\gamma_j = \Gamma(\mathbf{A}_j)$, $i = 1, \dots, n$ are the weights and $\delta_{\mathbf{s}_j}$'s are the point masses. When $d = 2$, take $\mathbf{s}_j = \left(\cos \left(\frac{2\pi(j-1)}{n} \right), \sin \left(\frac{2\pi(j-1)}{n} \right) \right) \in \mathbb{S}_d$, and arcs $\mathbf{A}_j = \left(\frac{2\pi(j - (\frac{3}{2}))}{n}, \frac{2\pi(j - (\frac{1}{2}))}{n} \right]$, $j = 1, \dots, n$. In higher dimensions, the \mathbf{A}_j 's are patches that partition the sphere \mathbb{S}_d with some *center* \mathbf{s}_j . Each of the coordinates of the $\vec{\gamma} = [\gamma_1, \dots, \gamma_n]'$ is an approximation to the mass of the patch containing \mathbf{s}_j , $j = 1, \dots, n$, in this case. The main principle behind the estimation of Γ is very simplistic: Given some grid $\mathbf{t}_j = \mathbf{s}_j$, $j = 1, \dots, n$ and either estimate ($\vec{\mathbf{I}}_{ECF}$ or $\vec{\mathbf{I}}_{PROJ}$) of $\vec{\mathbf{I}}$, invert the equation of $\vec{\mathbf{I}}$ above to get $\vec{\gamma}$. Using these weights and the grid $\mathbf{s}_1, \dots, \mathbf{s}_n$, $\hat{\Gamma}$ is defined by equation (4.8). The above method is formally correct. However, there are several numerical problems (Nolan, 2008).

4.4 Multivariate stable densities

The following theorem on the tails of the distribution of \mathbf{Y} due to Araujo & Giné (1980), provides an important result in understanding the relation between the spectral measure Γ and the distribution of \mathbf{Y} .

Theorem 4.5.1 (Nolan, 2014)

For a set $\mathbb{A} \in \mathbb{S}_d$, we define a cone generated by \mathbb{A} to be a $Cone(\mathbb{A}) = \{\mathbf{y} \in \mathbb{R}^d : |\mathbf{y}| > 0, \frac{\mathbf{y}}{|\mathbf{y}|} \in \mathbb{A}\} = \{r\mathbf{a} : r > 0, \mathbf{a} \in \mathbb{A}\}$.

$$\lim_{r \rightarrow \infty} \frac{P(\mathbf{Y} \in Cone(\mathbb{A}), |\mathbf{Y}| > r)}{P(|\mathbf{Y}| > r)} = \frac{\Gamma(\mathbb{A})}{\Gamma(\mathbb{S}_d)}$$

The mass that the spectral measure Γ assigns to \mathbb{A} determines the tail behavior of \mathbf{Y} in the direction of \mathbb{A} . In contrast, the local behavior is different from the directional tail behavior described in the theorem above (Abdul-Hamid, 1996).

Numerical methods are used to understand multivariate stable densities.

The inversion formula for the characteristic functions show that:

$$\begin{aligned} p(\mathbf{y}) &= (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-\mathbf{t}\langle \mathbf{y}, \mathbf{t} \rangle} e^{I_{\mathbf{Y}}(\mathbf{t})} d\mathbf{t} = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-\mathbf{t}\langle \mathbf{y}, \mathbf{t} \rangle + \Im I_{\mathbf{Y}}(\mathbf{t}) - \Re I_{\mathbf{Y}}(\mathbf{t})} d\mathbf{t} \\ &= \int_{\mathbb{R}^d} \cos[|\langle \mathbf{y}, \mathbf{t} \rangle| + \Im I_{\mathbf{Y}}(\mathbf{t})] e^{\Re I_{\mathbf{Y}}(\mathbf{t})} d\mathbf{t} \end{aligned}$$

Multivariate distributions are a very large class of distributions and cannot be parameterized by finite parameters. The use of multivariate stable laws in applications requires one to restrict the type of spectral measure. We describe some accessible classes:

- (1.) Independent components: The spectral measure is concentrated on the points where the coordinate axes intersect the sphere. Independence of the components makes it relatively easy to simulate densities and distribution functions in the univariate case.
- (2.) Discrete spectral measures: Λ , the spectral measure is discrete with point mass λ_j at locations s_j . Byczkowski et al. (1993) showed that this is a dense class such that for any spectral measure λ_1 , there exists a discrete measure λ_2 with a finite number of point masses where $|f_1(\mathbf{y}) - f_2(\mathbf{y})|$; the difference in the density functions

is uniformly small for all \mathbf{y} .

(3.) Elliptical contours: Here the joint characteristic function is of the form

$$\phi(\mathbf{u}) = \exp \left((\mathbf{u}^T Q \mathbf{u})^{\frac{\alpha}{2}} + i \mathbf{u} \cdot \delta \right)$$

where Q is a $d \times d$ positive definite shape matrix and δ is the shift vector. This class has an advantage in that it is computationally accessible and joint dependence is characterized by the set of pairwise parameters where $\frac{d(d-1)}{2}$ values are needed as in the Gaussian case.

Some basic properties of multivariate stable laws:

- Sums of independent stable random vectors are stable; the univariate projections $\mathbf{u} \cdot Y = \sum_k u_k Y_k$ are univariate stable laws.
- The support of stable laws are generally the entire space, however, in the one-dimensional case there are exceptions when $\alpha < 1$ and the spectral measure is one-sided.
- To be jointly stable, there has to be an α for which every component is univariate α -stable. Joint distributions can be constructed using copulas and vines. If multiple components have a similar index of stability(α), then it is logical to use a jointly stable model for those components and thereafter construct higher dimensional distributions. Unfortunately, full distributions are generally not jointly stable (Nolan, 2014).

4.5 Elliptically contoured stable distributions

4.5.1 Elliptically contoured stable laws

It can be shown that one can compute densities, make approximations on cumulative probabilities and fit elliptical stable distributions in dimensions $d \leq 40$.

If \mathbf{Y} is α -stable and elliptically contoured, then the joint characteristic function can be described as

$$E \left[e^{i\mathbf{u}^T \mathbf{Y}} \right] = \exp \left(-(\mathbf{u}^T \Sigma \mathbf{u})^{\frac{\alpha}{2}} + i\mathbf{u}^T \delta \right) \quad (4.12)$$

where the projection parameter functions are

$$\gamma(\mathbf{u}) = (\mathbf{u}^T \Sigma \mathbf{u})^{\frac{1}{2}}, \quad \beta(\mathbf{u}) = 0, \quad \delta(\mathbf{u}) = \langle \mathbf{u}, \delta \rangle$$

for some positive definite matrix Σ and shift vector $\delta \in \mathbb{R}^d$. We note that $\mathbf{y}^T \mathbf{z} = \sum_{k=1}^d y_k z_k$ is the inner product in \mathbb{R}^d . The spectral measure is complicated. However, it is a known measure. The matrix Σ is referred to as the *shape matrix* of the elliptical distribution.

Throughout this section, it is assumed that \mathbf{Y} is non-singular, which is equivalent to Σ being positive definite, that is, $\mathbf{u} \neq 0, \mathbf{u}^T \Sigma \mathbf{u} > 0$. All elliptically contoured stable distributions are scale mixtures of multivariate Normal distributions. Let $\mathbf{G} \sim N(\mathbf{0}, \Sigma)$ be a d -dimensional multivariate Normal random vector and $A \sim S(\frac{\alpha}{2}, 1, \gamma, 0)$ be an independent univariate positive $(\alpha/2)$ -stable random variable with $0 < \alpha < 2$. Then $\mathbf{Y} = A^{1/2} \mathbf{G}$ is an α -stable elliptically contoured distribution with joint characteristic function

$$\exp \left(- \left(\frac{\gamma}{2} \right)^{\alpha/2} \left(\sec \left(\frac{\pi \alpha}{4} \right) \right) (\mathbf{u}^T \Sigma \mathbf{u})^{\alpha/2} \right)$$

Elliptically contoured stable distributions are described as sub-Gaussian stable. Below, we describe a formula for simulating elliptical stable distributions when $0 < \alpha < 2, A \sim S(\frac{\alpha}{2}, 1, 2\gamma_0^2(\cos(\frac{\pi \alpha}{4})^{2/\alpha}), 0)$ where $\mathbf{G} \sim N(\mathbf{0}, \Sigma)$. Then,

$$\mathbf{Y} = A^{1/2} \mathbf{G} + \delta$$

has a characteristic function as in equation (4.18).

4.5.2 Densities of elliptically contoured stable laws

Let \mathbf{Y} be a d -dimensional α -stable elliptically contoured random vector with Σ (shape matrix) and δ (shift vector). Then $\mathbf{Y} \stackrel{d}{=} A^{1/2}\mathbf{G} + \delta$, where $A \sim S(\frac{\alpha}{2}, 1, 2\gamma_0^2(\cos(\frac{\pi\alpha}{4}))^{2/\alpha}, 0)$ and $\mathbf{G} \sim N(\mathbf{0}, \Sigma)$. Here, it is known that $\mathbf{G} \stackrel{d}{=} \Sigma^{1/2}\mathbf{G}_1$, where $\Sigma^{1/2}$ is derived from the Cholesky decomposition of Σ and $\mathbf{G}_1 \sim N(\mathbf{0}, \mathbf{I})$, \mathbf{G}_1 has independent standard Normal components. Therefore $\mathbf{G}_1 \stackrel{d}{=} A^{1/2}\Sigma^{1/2}\mathbf{G}_1 + \delta = \Sigma^{1/2}A^{1/2}\mathbf{G}_1 + \delta = \Sigma^{1/2}\mathbf{Y} + \delta$, where \mathbf{Y} is radially symmetric α -stable. So \mathbf{X} is an affine transformation of \mathbf{Y} and equation (4.17) shows that

$$f_{\mathbf{X}}(\mathbf{x}) = |\det \Sigma|^{-1/2} f_{\mathbf{Y}}(\Sigma^{-1/2}\mathbf{x}) = |\det \Sigma|^{-1/2} h\left(\Sigma^{-1/2}|\mathbf{x} - \delta|, \alpha, d\right) \quad (4.13)$$

(Nolan, 2006).

4.5.3 Statistical analysis as elliptical stable

Firstly, ways of assessing d -dimensional data set is described to see if it is approximately sub-Gaussian and then the parameters of a sub-Gaussian vector will be estimated.

A one-dimensional stable fit is carried out to each coordinate of the data using one of the univariate methods to obtain the estimates $\hat{\theta}_i = (\hat{\alpha}_i, \hat{\beta}_i, \hat{\gamma}_i, \hat{\delta}_i)$. If there are significant differences in the α_i 's, then the data cannot be described as jointly α -stable. Hence it is also not sub-Gaussian. Similarly, if the β_i 's are not all close to 0, then the distribution is asymmetric and cannot be sub-Gaussian. If all the α_i 's are close, they form a pooled estimate of $\alpha = \frac{(\sum_{i=1}^d \alpha_i)}{d}$ (the average of the indices of each component). Then, the data should be shifted by $\hat{\delta} = (\hat{\delta}_1, \hat{\delta}_2, \dots, \hat{\delta}_d)$. So, the distribution is centered at the origin.

The following step requires a test for sub-Gaussian behavior. We can approach this by analyzing two dimensional projections due to the following result:

If \mathbf{Y} is a d -dimensional sub-Gaussian α -stable random vector, then every two dimensional projection

$$\mathbf{Y} = (Y_1, Y_2) = (\mathbf{a}_1 \cdot \mathbf{Z}, \mathbf{a}_2 \cdot \mathbf{Z}) \quad (4.14)$$

where $(\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^d)$ is a two dimensional sub-Gaussian α -stable random vector. On the other hand, suppose \mathbf{Y} is a d -dimensional α -stable random vector with the property

that every two-dimensional projection of the form (4.34) is non-singular sub-Gaussian then d -dimensional \mathbf{Z} is non-singular sub-Gaussian α -stable. This leads to the assessment of multivariate data by looking at two dimensional distributions. This cannot be done for all projections. However, one can check pairs visually by glancing at scatterplots or by plotting the bivariate estimated directional scale function of the projected data and make comparisons to the scale function of the estimated elliptical fit.

There are two ways to estimate the $\frac{d(d+1)}{2}$ parameters. The upper triangular part of R . In the first method, we set $r_{\mathbf{u}} = \gamma_i^2$, that is, the square of the scale parameter in the i -th coordinate. Then, estimate r_{ij} by analyzing the pair (X_i, X_j) and take $r_{ij} = \frac{(\gamma^2(1,1) - r_{ii} - r_{jj})}{2}$ where $\gamma(1,1)$ is the scale parameter of $\langle(1,1), (Y_i, Y_j)\rangle = Y_i + Y_j$. This includes estimating $\frac{d + d(d-1)}{2} = \frac{d(d+1)}{2}$ one-dimensional scale parameters. In the second method, note that $E[e^{i\langle \mathbf{u}, \mathbf{Y} \rangle}] = e^{-((\mathbf{u}R\mathbf{u}^T)^{\frac{\alpha}{2}})}$ so

$$[\ln E[e^{i\langle \mathbf{u}, \mathbf{Y} \rangle}]]^{\frac{2}{\alpha}} = \mathbf{u}R\mathbf{u}^T = \sum_i u_i^2 r_{ii} + 2 \sum_{i < j} u_i u_j r_{ij}.$$
 The above equation is a linear function of the r'_{ij} s and, therefore, can be estimated by regression. This method is more accurate as it uses multiple directions whereas the first method uses only three directions: $(1,0)$, $(0,1)$ and $(1,1)$. Notice that $\mathbf{u}R\mathbf{u}^T = \gamma^2(\mathbf{u})$ is the square of the scale parameter in the direction \mathbf{u} . For $\gamma^2(\mathbf{u})$, sample estimates on a grid of \mathbf{u} points can be used for the middle term in the equation above. One should make sure that for both methods, the resulting matrix R should be positive definite (Nolan, 2005).

Chapter 5

Volatility model: GARCH(1,1)

The volatility clustering phenomenon is evident in this study where extreme returns cluster. We specify the conditional variance (volatility) by the GARCH(1,1) model. This is because the GARCH(1,1) model is considered to be a parsimonious model of conditional variance that fits many economic time series (Embrechts et al., 1999).

5.1 The ARCH model

The Autoregressive Conditional Heteroskedasticity (ARCH) model introduced by Engle (1982) has been extensively investigated by many researchers. Consider a log-return series w_t as

$$\begin{aligned}w_t &= \mu_t + a_t \\ a_t &= \sqrt{\sigma_t} \epsilon_t\end{aligned}\tag{5.1}$$

where ϵ_t is a white noise, $\epsilon_t \sim N(0, 1)$. The ARCH(m) process proposed by Engle (1982),

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2\tag{5.2}$$

where $\alpha_0 > 0, \alpha_i \geq 0$ are considered to ensure strictly positive variance. Generally, m is of high order because of the prominent volatility clustering phenomenon in financial markets. The unconditional variance is given by,

$$E[a_t^2] = \frac{\alpha_0}{1 - \sum_{i=1}^{max(m,s)} \alpha_i}\tag{5.3}$$

The process is covariance stationary if and only if the sum of the autoregressive parameters, $\sum_{i=1}^m \alpha_i < 1$ (Poon, 2008).

5.2 The GARCH model

An extension of the ARCH model is the generalized ARCH (referred to as GARCH) model. For a high order ARCH(m) process, it is more parsimonious to model volatility as a GARCH(m, s) due to Bollerslev (1986). For a log-return series w_t , we let $a_t = w_t - \mu_t$ be the innovation at time t . Hence, a_t follows a GARCH(m, s) model if

$$\begin{aligned} a_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \end{aligned} \quad (5.4)$$

where ϵ_t is a sequence of i.i.d. random variables with mean 0 and variance 1. We note that $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0$ and $\sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_j) < 1$. Here, it is also understood that $\alpha_i = 0$ for $i > m$ and $\beta_j = 0$ for $j > s$. The constraint $\alpha_i + \beta_i$ implies that the unconditional variance of a_t is finite whereas the conditional variance σ_t^2 evolves over time. Equation (5.1) above reduces to an ARCH(m) model if $s = 0$. The α_i and β_j are referred to as ARCH and GARCH parameters, respectively. For the properties of the GARCH model, the following representation is used. Let $\eta_t = a_t^2 - \sigma_t^2$ so $\sigma_t^2 = a_t^2 - \eta_t$. By substituting $\sigma_{t-i}^2 = a_{t-i}^2 - \eta_{t-i}$ ($i = 0, \dots, s$) into equation (5.1), we now write the GARCH model as

$$a_t^2 = \alpha_0 + \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) a_{t-i}^2 + \eta_t - \sum_{j=1}^s \beta_j \eta_{t-j} \quad (5.5)$$

where $E(\eta_t) = 0$ and $\text{Cov}(\eta_t, \eta_{t-j}) = 0$ for $j \geq 1$. Equation (5.2) is an ARMA form for the squared series a_t^2 . Therefore, the GARCH model resembles an ARMA model but with squared series a_t^2 . By using the unconditional mean of the ARMA model, we have

$$E[a_t^2] = \frac{\alpha_0}{1 - \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i)} \quad (5.6)$$

where the denominator is positive.

We now look at the simplest form of GARCH models, the GARCH(1,1) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad 0 \leq \alpha_1, \beta_1 \leq 1 \quad (\alpha_1 + \beta_1) < 1. \quad (5.7)$$

A large a_{t-1}^2 or σ_{t-1}^2 give rise to a large σ_t^2 . This means that a large a_{t-1}^2 tends to be followed by another large a_t^2 generating volatility clustering. It can be shown that if $1 - 2\alpha_1^2 - (\alpha_1 + \beta_1)^2 > 0$, then

$$\frac{E(a_t^4)}{[E(a_t^2)]^2} = \frac{3[1 - (\alpha_1 + \beta_1)^2]}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} > 3 \quad (5.8)$$

Similar to ARCH models, the tail distribution of the GARCH(1,1) model is heavier than the Normal distribution. For the GARCH(1,1) model, we assume that the forecast origin is h . The one-step ahead forecast is

$$\sigma_{h+1}^2 = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2$$

where a_h and σ_h^2 are known at time index h . The one-step ahead forecast is:

$$\sigma_h^2(1) = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2$$

$a_t^2 = \sigma_t^2 \epsilon_t^2$ is used for multi-step forecasts. Thus, the volatility equation is rewritten as

$$\sigma_{t+1}^2 = \alpha_0 + (\alpha_1 + \beta_1) \sigma_t^2 + \alpha_1 \sigma_t^2 (\epsilon_t^2 - 1)$$

When $t = h + 1$, we have

$$\sigma_{h+1}^2 = \alpha_0 + (\alpha_1 + \beta_1) \sigma_{h+1}^2 + \alpha_1 \sigma_{h+1}^2 (\epsilon_{h+1}^2 - 1)$$

Since $E(\epsilon_{h+1}^2 - 1 | \mathfrak{F}_h) = 0$, the two-step ahead volatility forecast at the origin satisfies the equation

$$\sigma_h^2(2) = \alpha_0 + (\alpha_1 + \beta_1) \sigma_h^2(1)$$

In general, we have

$$\sigma_h^2(\ell) = \alpha_0 + (\alpha_1 + \beta_1)\sigma_h^2(\ell - 1), \quad \ell > 1 \quad (5.9)$$

From equation (5.6), repeated substitutions show that the ℓ -step ahead forecast is:

$$\sigma_h^2(\ell) = \frac{\alpha_0 [1 - (\alpha_1 + \beta_1)^2]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{\ell-1} \sigma_h^2(1),$$

Hence,

$$\sigma_h^2(\ell) \rightarrow \frac{\alpha_0}{1 - \alpha_1 - \beta_1}, \quad \ell \rightarrow \infty$$

only if $\alpha_1 + \beta_1 < 1$. With the GARCH(1,1) model, the multistep ahead volatility forecasts converge to the unconditional variance of a_t as the forecast horizon tends to infinity if $\text{Var}(a_t)$ exists (Tsay, 2005).

5.3 Parameter estimation

Yang (2012b) defines the GARCH(m, s) as

$$\begin{aligned} \sigma_t &= \alpha_0 + \sum_{i=1}^m \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^s \beta_i \sigma_{t-i} \\ \epsilon_t &= v_t \sqrt{\sigma_t} \end{aligned}$$

where σ_t defines the conditional variance and v_t is the white noise term

To estimate parameters of GARCH models with given k, m and s , we have

$$y_t = C + \sum_{i=1}^k a_i y_{t-i} + \epsilon_t \quad (5.10)$$

$$\epsilon_t = v_t \sqrt{\sigma_t} \quad (5.11)$$

$$\sigma_t = \alpha_0 + \sum_{i=1}^m \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^s \beta_i \sigma_{t-i} \quad (5.12)$$

where v_t is the white noise term. ϵ_t follows a Normal distribution with mean zero and conditional variance σ_t , i.e.

$$p(\epsilon_t | \epsilon_{t-1}, \dots, \epsilon_0) = \frac{1}{\sqrt{2\pi\sigma_t}} \exp\left(-\frac{\epsilon_t^2}{2\sigma_t}\right) \quad (5.13)$$

The log-likelihood function of the parameter vector $\theta = (\alpha_0, \alpha_1, \dots, \alpha_s, \beta_1, \dots, \beta_m)^T$ is

$$L(\theta) = \sum_{t=s+1}^n l_t(\theta) = \sum_{t=s+1}^n \left(-\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \sigma_t - \frac{\epsilon_t^2}{2\sigma_t} \right) \quad (5.14)$$

We, therefore, have

$$\frac{\partial l_t(\theta)}{\partial \theta} = \left(\frac{\epsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t} \right) \frac{\partial \sigma_t}{\partial \theta} \quad (5.15)$$

$$\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta^T} = \left(\frac{\epsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t} \right) \frac{\partial^2 \sigma_t}{\partial \theta \partial \theta^T} + \left(\frac{1}{2\sigma_t^2} - \frac{\epsilon_t^2}{\sigma_t^3} \right) \frac{\partial \sigma_t}{\partial \theta} \frac{\partial \sigma_t}{\partial \theta^T} \quad (5.16)$$

where

$$\frac{\partial \sigma_t}{\partial \theta} = (1, \epsilon_{t-1}^2, \dots, \epsilon_{t-s}^2, \sigma_{t-1}, \dots, \sigma_{t-m})^T + \sum_{i=1}^m \beta_i \frac{\partial \sigma_{t-i}}{\partial \theta} \quad (5.17)$$

The gradient is

$$\nabla L(\theta) = \frac{1}{2} \sum_{t=s+1}^n \left(\frac{\epsilon_t^2}{\sigma_t^2} - \frac{1}{\sigma_t} \right) \frac{\partial \sigma_t}{\partial \theta} \quad (5.18)$$

and the Fisher information matrix is given by

$$\begin{aligned} J &= \sum_{t=s+1}^n E \left[\left(\frac{\epsilon_t^2}{2\sigma_t^2} - \frac{1}{2\sigma_t} \right) \frac{\partial^2 \sigma_t}{\partial \theta \partial \theta^T} + \left(\frac{1}{2\sigma_t^2} - \frac{\epsilon_t^2}{\sigma_t^3} \right) \frac{\partial \sigma_t}{\partial \theta} \frac{\partial \sigma_t}{\partial \theta^T} \right] \\ &= -\frac{1}{2} \sum_{t=s+1}^n E \left(\frac{1}{h_t^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta^T} \right) \end{aligned} \quad (5.19)$$

We now consider the GARCH(1,1) model

$$\epsilon_t = v_t \sqrt{\sigma_t} \quad (5.20)$$

$$\sigma_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (5.21)$$

to estimate the coefficients $\theta = (\alpha_0, \alpha_1, \beta_1)^T$, where

$$\nabla L(\theta) = \frac{1}{2} \sum_{t=2}^n \left(\frac{\epsilon_t^2}{\sigma_t^2} - \frac{1}{\sigma_t} \right) \frac{\partial \sigma_t}{\partial \theta} \quad (5.22)$$

and

$$J = -\frac{1}{2} \sum_{t=2}^n E \left(\frac{1}{\sigma^2} \frac{\partial \sigma_t}{\partial \theta} \frac{\partial \sigma_t}{\partial \theta^T} \right) \quad (5.23)$$

with

$$\frac{\partial \sigma_t}{\partial \theta} = (1, \epsilon_{t-1}^2, \sigma_{t-1})^T + \beta_1 \frac{\partial \sigma_{t-1}}{\partial \theta} \quad (5.24)$$

This chapter explored the properties and parameter estimation of GARCH models. GARCH models are popular in the financial industry due to the properties of capturing most of the stylized facts of financial data. In this research we investigate a case of the hybrid, stable-GARCH(1,1) model to see the effect of a GARCH filter on modeling BRICS exchange rates. The next chapter discusses the research methodology thereafter we look at the empirical results.

Chapter 6

Methodology

This chapter discusses the relevant research methodology.

6.1 Autocorrelation

6.1.1 Autocorrelation

Autocorrelation is the term used to define the correlation between the members of a series of observations ordered in time. In the classical linear regression model, it is assumed that the autocorrelation does not exist in the disturbances u_i . That is,

$$E(u_i u_j) = 0, \quad i \neq j \quad (6.1)$$

6.1.2 Autocorrelation function

For a stationary process Y_t with mean $E(Y_t) = \mu$ and variance $\text{Var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2$ which is constant. The covariance between Y_t and Y_{t+k} is defined as:

$$\gamma_k = \text{Cov}(Y_t, Y_{t+k}) = E(Y_t - \mu)(Y_{t+k} - \mu) \quad (6.2)$$

and the autocorrelation function (ACF) is defined as:

$$\rho_k = \frac{\text{Cov}(Y_t, Y_{t+k})}{\sqrt{\text{Var}(Y_t)}\sqrt{\text{Var}(Y_{t+k})}} = \frac{\gamma_k}{\gamma_0} \quad (6.3)$$

where k denotes a separation by k lags (Wei, 2006).

6.1.3 Partial autocorrelation function

One may want to investigate the correlation between Y_t and Y_{t+k} after the mutual dependence of the intervening variables $Y_{t+1}, Y_{t+2}, \dots, Y_{t+k-1}$ have been removed. The conditional correlation

$$\text{Corr}(Y_t, Y_{t+k} | Y_{t+1}, \dots, Y_{t+k-1}) \quad (6.4)$$

is referred to as partial correlation which for convenience is denoted by ϕ_{kk} . Consider a regression model which is regressed on k lagged variables, that is

$$Y_{t+k} = \phi_{k1}Y_{t+k-1} + \phi_{k2}Y_{t+k-2} + \dots \phi_{kk}Y_t + e_{t+k}$$

where ϕ_{ki} denotes the i th regression parameter and e_{t+k} is the error term. It can be shown, by using Cramer's rule, that the PACF, as denoted by Wei (2006):

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}} \quad (6.5)$$

6.1.4 Ljung-Box test for autocorrelation

Formally, the Ljung-Box test defines serial correlation in the alternative hypothesis where the test statistic is given by

$$\mathcal{Q} = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k}$$

where n is the sample size, the sample autocorrelation at lag k is $\hat{\rho}_k$. Under H_0 , the test statistic, \mathcal{Q} follows a chi-squared distribution with h degrees of freedom. Thus, we reject H_0 at α level of significance if the value of \mathcal{Q} exceeds the $(1 - \alpha)$ - quantile of the

chi-squared distribution with h degrees of freedom (Shumway & Stoffer, 2010).

6.2 Periodogram

Wei (2006) describes the periodogram as follows:

Given a time series of n observations, we can represent these n observations in the following Fourier representation:

$$Y_t = \sum_{k=0}^{\frac{n}{2}} (a_k \cos \omega_k t + b_k \sin \omega_k t) \quad (6.6)$$

where $\omega_k = \frac{2\pi k}{n}$, $k = 0, 1, \dots, \frac{n}{2}$ are Fourier frequencies. Also,

$$a_k = \begin{cases} \frac{1}{n} \sum_{t=1}^n Y_t \cos \omega_k t; & k = 0 \text{ and } k = \frac{n}{2} \text{ if } n \text{ is even} \\ \frac{2}{n} \sum_{t=1}^n Y_t \cos \omega_k t, & k = 1, 2, \dots, \frac{n-1}{2} \end{cases} \quad (6.7)$$

and

$$b_k = \frac{2}{n} \sum_{t=1}^n Y_t \sin \omega_k t, \quad k = 1, 2, \dots, \frac{n-1}{2} \quad (6.8)$$

are Fourier coefficients. These coefficients are the least squares estimates of the coefficients in fitting the regression model below:

$$Y_t = \sum_{k=0}^{\frac{n}{2}} (a_k \cos \omega_k t + b_k \sin \omega_k t) + \epsilon_t \quad (6.9)$$

The quantity $I(\omega_k)$ is defined as:

$$I(\omega_k) = \begin{cases} na_0^2, & k=0 \\ \frac{n}{2}(a_k^2 + b_k^2), & k = 1, 2, \dots, \frac{n-1}{2} \\ na_{\frac{n}{2}}^2, & k = \frac{n}{2} \text{ when } n \text{ is even} \end{cases} \quad (6.10)$$

$I(\omega_k)$ is called the periodogram which was introduced by Schuster (1898) to search for periodic components in a series.

6.2.1 Testing for hidden periodicity

If it is believed that a time series contains a periodic component, the underlying frequency is most likely to be unknown. For the model

$$Y_t = \mu + \alpha \cos \omega t + \beta \sin \omega t + e_t$$

we test the hypotheses

$$H_0 : \alpha = \beta = 0 \text{ versus } H_1 : \alpha \neq 0 \text{ or } \beta \neq 0$$

where e_t is a Gaussian white noise series with i.i.d $N(0, \sigma^2)$ and frequency ω is not known. If the time series contains a periodic component at frequency ω , then we assume that the periodogram $I(\omega_k)$ at a Fourier frequency ω_k closest to ω will be the maximum. We can identify the maximum periodogram ordinate and, thereafter, test if this ordinate is considered to be maximum in a random sample of $\frac{n}{2}$ i.i.d. random variables. Each random variable is a multiple of a chi-squared distribution with two degrees of freedom. Here, the test statistic is

$$I^{(1)}(\omega_{(1)}) = \max\{I(\omega_k)\} \quad (6.11)$$

where $\omega_{(1)}$ indicates the Fourier frequency with the maximum periodogram ordinate. The null hypothesis H_0 is that the series Y_t is a Gaussian white noise $N(0, \sigma^2)$. Therefore, the periodogram ordinates $\frac{I(\omega_k)}{\sigma^2}, k = 1, 2, \dots, \frac{n}{2}$ are i.i.d chi-squared distributed random variables with two degrees of freedom and have density

$$p(x) = \frac{1}{2} \exp\left(-\frac{x}{2}\right), \quad 0 \leq x \leq \infty. \quad (6.12)$$

Then, for any $g \geq 0$, we have

$$\begin{aligned} P\left[\frac{I^{(1)}(\omega_{(1)})}{\sigma^2} > g\right] &= 1 - P\left[\frac{I^{(1)}(\omega_{(1)})}{\sigma^2} \leq g\right] \\ &= 1 - P\left\{\frac{I^{(1)}(\omega_{(1)})}{\sigma^2} \leq g, \quad k = 1, 2, \dots, \frac{n}{2}\right\} \\ &= 1 - \left\{\int_0^g e^{-\frac{x}{2}} dx\right\}^{\frac{n}{2}} \end{aligned}$$

$$= 1 - (1 - e^{-\frac{g}{2}})^{\frac{n}{2}}$$

If σ^2 were known then the above equation would have been used to derive an exact test for the maximum ordinate. However, in practice σ^2 is unknown and is required to be estimated. In order to derive an unbiased estimator, we should take note that under H_0 ,

$$E \left[\sum_{k=1}^{\frac{n}{2}} I(\omega_k) \right] = \left[\frac{n}{2} \right] 2\sigma^2$$

Hence,

$$\hat{\sigma}^2 = \frac{1}{2^{\frac{n}{2}}} \sum_{k=1}^{\frac{n}{2}} I(\omega_k) \quad (6.13)$$

is an unbiased estimator of σ^2 . This leads to the test statistic

$$V = \frac{I^{(1)}(\omega_{(1)})}{2^{\frac{n}{2}}} \sum_{k=1}^{\frac{n}{2}} I(\omega_k). \quad (6.14)$$

Using the fact that $I(\omega_k), k = 1, 2, \dots, \frac{n}{2}$ are independent, we have

$$\text{Var}(\sigma^2) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

It also follows that $\hat{\sigma}^2$ is a consistent estimator of σ^2 . So for large samples, the distribution of V can be approximated by the same distribution as $\frac{I^{(1)}(\omega_{(1)})}{\sigma^2}$ for any $g \geq 0$,

$$P(V > g) \simeq 1 - (1 - e^{-\frac{g}{2}})^{\frac{n}{2}}. \quad (6.15)$$

The exact test for $\max\{I(\omega_k)\}$ was derived by Fisher (1929) based on the statistic

$$T = \frac{I^{(1)}(\omega_{(1)})}{\sum_{k=1}^{\frac{n}{2}} I(\omega_k)} \quad (6.16)$$

The null hypothesis H_0 which states that the series Y_t is a Gaussian white noise process.

Fisher (1929) also showed that

$$P(T > g) = \sum_{j=1}^m (-1)^{(j-1)} \binom{N}{j} (1 - jg)^{N-1} \quad (6.17)$$

where $N = \frac{n}{2}, g > 0$ and m is the largest integer less than $\frac{1}{g}$. For any given significance level α , equation (5.18) can be used to find critical value g_α such that,

$$P(T > g_\alpha) = \alpha$$

We reject the null hypothesis if the T value calculated from the series is larger than g_α and conclude that the series Y_t contains a periodic component. This test procedure is known as Fisher's test, also known as Fisher's Kappa (Wei, 2006).

6.3 Tests for stationarity

6.3.1 The unit root test

We start with the unit root stochastic process:

$$Y_t = \rho Y_{t-1} + u_t, \quad -1 \leq \rho \leq 1 \quad (6.18)$$

where u_t is a white noise error term. If $\rho = 1$, in the case of the unit root, the equation above becomes a random walk model without drift which is known to be a non-stationary stochastic process. We regress Y_t on its lagged value Y_{t-1} to determine if the estimated ρ is statistically equal to 1. If so, then Y_t is considered to be non-stationary. This is the general idea behind unit root tests for stationarity. The equation above is manipulated by subtracting Y_{t-1} from both sides to obtain

$$\begin{aligned} Y_t - Y_{t-1} &= \rho Y_{t-1} - Y_{t-1} + u_t \\ &= (\rho - 1)Y_{t-1} + u_t \end{aligned}$$

This can be alternatively written as

$$\Delta Y_t = \delta Y_{t-1} + u_t \quad (6.19)$$

where $\delta = (\rho - 1)$ and Δ is the first difference operator. We estimate equation (6.19) and test the null hypothesis

$$H_0 : \delta = 0$$

If $\delta = 0$, then $\rho = 1$. This implies a unit root which, in turn, means the time series under consideration is non-stationary. Note that if $\delta = 0$, then equation (6.19) becomes

$$\Delta Y_t = (Y_t - Y_{t-1}) = u_t \quad (6.20)$$

where u_t is the white noise error term and is stationary, which means that the first differences of a random walk time series are stationary. To estimate equation (6.20), we take the first differences of Y_t and regress them on Y_{t-1} and evaluate the estimated slope coefficient in this regression ($= \hat{\delta}$) is equal to zero or otherwise. If it is zero, we

conclude that Y_t is non-stationary. However, if it is negative, we may conclude that Y_t to be stationary.

There are several decisions to be considered in the implementation of the DF test procedure. Note that a random walk process may have no drift, it may have drift or there may be both deterministic and stochastic trends.

The DF test is estimated under three different null hypotheses:

1. Y_t is a random walk:
$$\Delta Y_t = \delta Y_{t-1} + u_t$$
2. Y_t is a random walk with drift:
$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + u_t$$
3. Y_t is a random walk with drift around a stochastic trend:
$$\Delta Y_t = \beta_1 + \beta_2 t \delta Y_{t-1} + u_t,$$

where t is the time or trend variable. In each of the above cases, we have $H_0 : \delta = 0$, that is, there exists a unit root and the time series is non-stationary. The alternative hypothesis states that: $H_1 : \delta < 0$ which implies the time series is stationary. If H_0 is rejected, then Y_t is a stationary time series with zero mean in the first case. In the second case, it is implied that Y_t is stationary with non-zero mean $\left(\frac{\beta_1}{(1-\rho)} \right)$ and lastly, Y_t is stationary around a deterministic trend. The actual estimation is done by an ordinary least squares method where we divide the estimated coefficient of Y_t in each case by its standard error to compute the τ statistic. With reference to the DF tables or any other statistical package, if the computed absolute value of the τ statistic exceeds the DF critical τ values, we fail to reject the null hypothesis, in which case the time series is non-stationary (Gujarati, 2004).

6.3.2 Augmented Dickey-Fuller test

In the DF test described in the previous section, it was assumed that the error term u_t is uncorrelated. Conversely, in the case where the error terms are correlated, Dickey and Fuller devised another test known as the Augmented Dickey-Fuller (ADF)

test. The ADF test is conducted by augmenting the equations above by adding the lagged values of the dependent variable ΔY_t . We are concerned in estimating the following regression equation

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + \epsilon_t \quad (6.21)$$

where ϵ_t is a pure white noise error term and where $\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2})$, $\Delta Y_{t-2} = (Y_{t-2} - Y_{t-3})$ and so forth. The number of lagged differences to be added is determined empirically where the main idea is to include enough terms so that the error term is serially uncorrelated. In the ADF test, we test $\delta = 0$. The same critical values are used since the ADF test follows the same asymptotic distribution as the DF statistic (Gujarati, 2004).

6.3.3 Philips-Perron test

The Phillips-Perron (P-P) test a more comprehensive method to test unit root non-stationarity. This test is similar to the ADF test, but includes an automatic correction to the DF test to allow for autocorrelated residuals.

The Phillips-Perron test fits the following regression

$$Y_t = \alpha + \rho Y_{t-1} + \epsilon_t \quad (6.22)$$

where a constant term may be excluded or a trend term may be included. Z_ρ and Z_τ are two statistics that are calculated as follows:

$$Z_\rho = n(\hat{\rho}_n - 1) - \frac{1}{2} \frac{n^2 \hat{\sigma}^2}{s_n^2} (\hat{\lambda}_n^2 - \hat{\gamma}_{0,n}) \quad (6.23)$$

$$Z_\tau = \sqrt{\frac{\hat{\gamma}_{0,n}}{\hat{\lambda}_n^2}} \frac{\hat{\rho}_n - 1}{\hat{\sigma}} - \frac{1}{2} (\hat{\lambda}_n^2 - \hat{\gamma}_{0,n}) \frac{1}{\hat{\lambda}_n} \frac{n \hat{\sigma}}{s_n} \quad (6.24)$$

$$\hat{\gamma}_{j,n} = \frac{1}{n} \sum_{t=j+1}^n \hat{u}_t \hat{u}_{t-j}$$

$$\hat{\lambda}_n^2 = \hat{\lambda}_{0,n} + 2 \sum_{j=1}^q \left(1 - \frac{j}{q+1}\right) \hat{\lambda}_{j,n}$$

$$s_n^2 = \frac{1}{n-k} \sum_{t=1}^m \hat{u}_t^2$$

where u_t is the OLS residual, k is the number of covariates in the regression model, q is the number of lags to use in calculating $\hat{\lambda}_n^2$ and $\hat{\sigma}$ is the OLS standard error of $\hat{\rho}$.

The regression is Y on lagged Y and not the differenced Y on lagged Y . Z_τ is the adjusted t -statistic as in the DF test:

$$\frac{\hat{\rho}_n - 1}{\hat{\sigma}} \quad (6.25)$$

s_n^2 is an unbiased OLS estimator of the variance of the error terms.

$$\hat{\gamma}_{j,n} = \frac{1}{n} \sum_{t=j+1}^n \hat{u}_t \hat{u}_{t-j}$$

When $j = 0$, this is a maximum likelihood estimate of the variance of the error terms when the estimator of the covariance between two error terms is j periods apart.

$$\hat{\lambda}_n^2 = \hat{\gamma}_{0,n} + 2 \sum_{j=1}^q \left(1 - \frac{j}{q+1}\right) \hat{\gamma}_{j,n}$$

where q is the number of lagged covariances. When the covariances are zero; that is, when the autocorrelation between error terms $\hat{\gamma}_{j,n}$ is zero for $j > 0$. Thus, the second term in the equation is eliminated and $\hat{\lambda}_n^2 = \hat{\gamma}_{0,n}$. We can make a replacement for Z_τ which becomes

$$Z_\tau = \sqrt{\frac{\hat{\gamma}_{0,n}}{\hat{\lambda}_n^2} \frac{\hat{\rho}_n - 1}{\hat{\sigma}}} - \frac{1}{2} \left(\hat{\lambda}_n^2 - \hat{\gamma}_{0,n} \right) \frac{1}{\hat{\lambda}_n} \frac{n\hat{\sigma}}{s_n} \quad (6.26)$$

In this case $\hat{\lambda}_n^2 - \hat{\gamma}_{0,n} = 0$, and the second term vanishes. $\frac{\hat{\gamma}_{0,n}}{\hat{\lambda}_n^2} = 1$ hence the term $\sqrt{\frac{\hat{\gamma}_{0,n}}{\hat{\lambda}_n^2} \frac{\hat{\rho}_n - 1}{\hat{\sigma}}}$ reduces to $\frac{\hat{\rho}_n - 1}{\hat{\sigma}}$ and therefore $Z_\tau = \frac{\hat{\rho}_n - 1}{\hat{\sigma}}$. We notice the similarity as in the standard DF test. When there is no autocorrelation between the error terms, this term of the P-P test is equal to the DF test. The P-P test corrects the DF test for autocorrelation amongst error terms outside of a regression framework (non-parametrically). The critical values have the same distribution as the DF statistic.

If there is no autocorrelation between error terms, then the covariances are equal and the second term in the other P-P statistic becomes zero as $\hat{\lambda}_n^2 = \hat{\gamma}_{0,n}$

$$Z_\rho = n(\hat{\rho}_n - 1) - \frac{1}{2} \frac{n^2 \hat{\sigma}^2}{s_n^2} \left(\hat{\lambda}_n^2 - \hat{\gamma}_{0,n} \right) \quad (6.27)$$

In the above case $Z_\rho = n(\hat{\rho}_n - 1)$ which again is the same as the DF test (StataCorp, 2015).

6.3.4 Kwiatkowski-Phillips-Schmidt-Shin test

The Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) has a null hypothesis of stationarity in a time series around a mean or a linear trend versus the alternative hypothesis which states that the time series is non-stationary due to the presence of a unit root. The KPSS model is made up by a series of observations represented as a sum of three components, namely: a deterministic trend, a random walk and a stationary error term. The model is as follows:

$$Y_t = \xi t + r_t + \epsilon_t \quad (6.28)$$

$$r_t = r_{t-1} + u_t \quad (6.29)$$

where $Y_t, t = 1, 2, \dots, T$ denotes series of observations of the variable of interest, t is the deterministic trend, r_t is a random walk process and ϵ_t is the error term, which by assumption is stationary. u_t is the error term in the second equation above and is a series of i.i.d. with mean 0 and variance $\hat{\sigma}_u^2$. The null hypothesis of stationarity is equivalent to the assumption that $\hat{\sigma}_u^2$ of the random walk process r_t equals to zero. When $\xi = 0$, the null hypothesis implies Y_t is stationary around r_0 . Conversely, if $\xi \neq 0$, then this suggests Y_t is stationary around a linear trend. If $\hat{\sigma}_u^2 > 0$, then Y_t is non-stationary due to the presence of a unit root.

By subtracting Y_t from both sides of the first equation above, we have

$$\Delta Y_t = \xi + u_t + \Delta \epsilon_t = \xi + w_t \quad (6.30)$$

where w_t by the assumption that the error terms (ϵ_t and u_t) are i.i.d. is generated by an AR(1) process where: $w_t = v_t + \theta v_{t-1}$. The KPSS model may be described as:

$$Y_t = \xi + \beta Y_{t-1} + w_t \quad (6.31)$$

$$w_t = v_t + \theta v_{t-1}, \beta = 1 \quad (6.32)$$

The above equations show an interesting relationship between the KPSS and the DF test. In the DF test, $\beta = 1$ on the assumption that $\theta = 0$ where θ is the nuisance parameter. Kwiatkowski et al. (1992) made an assumption that β is the the nuisance parameter and, hence, tests if $\theta = -1$. Assuming $\beta = 0$, a one-side Lagrange Multiplier (LM) test with $H_0 : v\sigma_u^2 = 0$ where u_t is Normally distributed and ϵ_t are i.i.d.. random variables with a zero mean and constant variance σ_ϵ^2

The KPSS test statistics is as follows:

- (a) Testing a null hypothesis of stationarity around a linear trend versus the alternative hypothesis of the presence of a unit root

Let $e_t, t = 1, 2, 3, \dots, T$ denote the estimated errors from a regression on Y_t and $\hat{\sigma}_t^2$ denotes the estimated variance which is equal to the sum of error squares divided by the number of observations T . Partial sums of errors are computed from:

$$S_t = \sum_{t=1}^t e_t \text{ for } t = 1, 2, \dots, T$$

The LM test statistic is then defined as

$$\text{LM} = \sum_{t=1}^T \frac{S_t^2}{\sigma_\epsilon^2} \quad (6.33)$$

- (b) Testing the null hypothesis of stationarity around the mean versus an alternative hypothesis of the presence of a unit root

e_t , the estimated errors are computed as residuals of regression on Y_t ; that is $e_t = Y_t - \bar{Y}$ and the remaining definitions are left unchanged. The long-run variance is:

$$\sigma^2 = \lim T^{-1} E [S_T^2] \quad (6.34)$$

the long-run variance appears when we define an asymptotic distribution of a test statistic. Kwiatkowski et al. (1992) provide a consistent estimate of the long-run variance which is given by

$$s^2(k) = T^{-1} \sum_{t=1}^T e_t^2 + 2T^{-1} \sum_{j=1}^k w(j, k) \sum_{t=j+1}^T e_t e_{t-1} \quad (6.35)$$

where $w(j, k)$ denote weights depending on the choice of a spectral window. The Bartlett window is used by Kwiatkowski et al. (1992) where the weights $w(j, k) = 1 - \frac{j}{k+1}$. This makes certain that $s^2(k)$ is non-negative. It is argued that for quarterly data, lag $k = 8$ is the ideal choice as if $k < 8$, then the size of test is distorted and if $k > 8$ the power decreases. Here, the KPSS test statistic is computed as a ratio of sum of squared partial sums and the estimated long-run variance

$$\hat{\eta} = T^{-2} \frac{\sum S_t^2}{s^2(k)} \quad (6.36)$$

The symbols $\hat{\eta}_\mu$ and $\hat{\eta}_t$ denotes the KPSS test statistic for testing stationarity around a mean and around a trend, respectively. The asymptotic distribution of the KPSS test statistic is non-standard and converges to a Brownian bridge of higher order.

The $\hat{\eta}_\mu$ statistic for testing stationarity around mean converges to

$$\hat{\eta}_\mu \rightarrow \int_0^1 V(r)^2 dr$$

In the above equation, the standard Brownian bridge is denoted by: $V(r) = W(r) - rW(1)$ which is defined for a standard Wiener process $W(r)$.

The KPSS test statistic for $\hat{\eta}_t$ is testing for stationarity around a trend, when $\xi \neq 0$, weakly converges to a second order Brownian bridge, $V_2(r)$, which is defined as :

$$V_2(r)_2 = W(r) + (2r - 3r^2)W(1) + (-6r + 6r^2) \int_0^1 W(s) ds$$

The test statistic above converges weakly to the limit

$$\hat{\eta}_t \rightarrow \int_0^1 V_2(r)^2 dr$$

In summary, the KPSS test is performed as follows:

- (i) We test the hypotheses H_0 : stationarity around a mean or around a trend versus the alternative H_1 : non-stationarity of a time series due the presence of a unit root.
- (ii) Compute the value of the test statistic
- (iii) If the computed value is greater than the critical value at a given level of significance, the null hypothesis of stationarity is rejected (Syczewska et al., 2010).

6.4 Measures of dependency

6.4.1 Scatterplot matrices

If we have a set of variables Y_1, Y_2, \dots, Y_k , then the scatter plot matrix contains all the pairwise scatterplots of the variables in a matrix format. That is, if there are k variables, then the scatterplot matrix has k rows and k columns. The i th row and j th column of the matrix is the plot of Y_i versus Y_j (NIST/SEMATECH, 2013).

6.4.2 Covariance and correlation matrices

In order to measure the linear dynamic dependence of a stationary time series \mathbf{y}_t , we define its lag k cross-covariance matrix as

$$\begin{aligned}\mathbf{\Gamma}_k &= \text{Cov}(\mathbf{y}_t, \mathbf{y}_{t-k}) \\ &= E[(\mathbf{y}_t - \mu)(\mathbf{y}_{t-k} - \mu)'] \\ &= \begin{bmatrix} E(\tilde{y}_{1t}\tilde{y}_{1,t-k}) & E(\tilde{y}_{1t}\tilde{y}_{2,t-k}) & \cdots & E(\tilde{y}_{1t}\tilde{y}_{n,t-k}) \\ \vdots & \vdots & \cdots & \vdots \\ E(\tilde{y}_{nt}\tilde{y}_{1,t-k}) & E(\tilde{y}_{nt}\tilde{y}_{1,t-k}) & \cdots & E(\tilde{y}_{nt}\tilde{y}_{n,t-k}) \end{bmatrix} \end{aligned} \quad (6.37)$$

where $\mu = E(\mathbf{y}_t)$ is the mean vector of \mathbf{y}_t and $\tilde{\mathbf{y}}_t = (\tilde{y}_{1t}, \dots, \tilde{y}_{nt})' \equiv \mathbf{y}_t - \mu$ is the mean-adjusted time series. The above cross-covariance matrix is a function of k and not the time index t , since \mathbf{y}_t is stationary. If $k = 0$, we have the covariance matrix $\mathbf{\Gamma}_0$ of \mathbf{y}_t . The (i, j) th element of $\mathbf{\Gamma}_k$ is denoted as $\gamma_{k,ij}$. From the matrix above, we see that $\gamma_{k,ij}$ is the covariance between $y_{i,t}$ and $y_{j,t-k}$. Therefore, for the positive lag k , $\gamma_{k,ij}$ can be regarded as a measure of linear dependence of the i th component y_{it} on the k th lagged value of the component y_{jt} .

From the definition of covariance depicted by the matrix above, for a negative lag k :

$$\begin{aligned}
\mathbf{\Gamma}_k &= E \left[(\mathbf{y}_t - \mu)(\mathbf{y}_{t-k} - \mu)' \right] \\
&= E \left[(\mathbf{y}_{t+k} - \mu)(\mathbf{y}_t - \mu)' \right] \text{ (because of stationarity)} \\
&= \left\{ E \left[(\mathbf{y}_t - \mu)(\mathbf{y}_{t+k} - \mu')' \right] \right\}' \text{ since } \mathbf{A} = (\mathbf{A}')' \\
&= \left\{ E \left[(\mathbf{y}_t - \mu)(\mathbf{y}_{t-(-k)} - \mu')' \right] \right\}' \\
&= \left\{ \mathbf{\Gamma}_{-k} \right\}' \text{ (by definition)} \\
&= \mathbf{\Gamma}_{-k}'
\end{aligned}$$

Unlike the case of the univariate stationary time series, where the autocovariances of lag k and lag $-k$ are identical, one needs to take the transpose of a positive lag cross-covariance matrix to obtain the negative lag cross-covariance matrix.

For a stationary multivariate linear time series \mathbf{y}_t , we have (for $k \geq 0$),

$$\begin{aligned}
\mathbf{\Gamma}_k &= E \left[(\mathbf{y}_t - \mu)(\mathbf{y}_{t-k} - \mu)' \right] \\
&= E \left[(\mathbf{a}_t + \psi_1 \mathbf{a}_{t-1} + \cdots)(\mathbf{a}_{t-k} + \psi_1 \mathbf{a}_{t-k-1} + \cdots)' \right] \\
&= E \left[(\mathbf{a}_t + \psi_1 \mathbf{a}_{t-1} + \cdots)(\mathbf{a}_{t-k}' + \psi_1' \mathbf{a}_{t-k-1}' + \cdots) \right] \\
&= \sum_{i=k}^{\infty} \psi_i \mathbf{\Sigma}_a \psi_{i-k}'
\end{aligned}$$

where the last inequality holds as \mathbf{a}_t has no serial correlations and $\psi_0 = \mathbf{I}_k$.

For a stationary series \mathbf{y}_t , the lag k cross-correlation matrix ρ_k is defined as

$$\rho_k = \mathbf{D}^{-1} \Gamma_k \mathbf{D}^{-1} = [\rho_{k,ij}] \quad (6.38)$$

where $\mathbf{D} = \text{diag}\{\sigma_1, \dots, \sigma_k\}$ is the diagonal matrix of standard deviations of the components of y_t . More specifically, $\sigma_i^2 = \text{Var}(y_{it}) = \gamma_{0,ii}$, i.e the (i, i) th element of Γ_0 . ρ_0 is symmetric with diagonal elements being 1. Off-diagonal elements of ρ_0 are correlations between components in \mathbf{y}_t . For $k > 0$, ρ_k is not symmetric since ρ_{ij} is the correlation coefficient between y_{it} and $y_{j,t-k}$ whereas $\rho_{k,ji}$ is the correlation coefficient between y_{jt} and $y_{i,t-k}$. Using the properties of Γ_k , we have, $\rho_k = \rho'_{-k}$ (Tsay, 2013).

6.5 Goodness-of-fit tests

There has been a long standing debate on whether it is required to fit stable distributions, especially when analyzing returns on financial data. The Normality assumption is the core of modern portfolio theory. This assumption cannot justify the characteristics of heavy-tails and skewness. Stable distributions are proposed as a better model for financial asset returns. The Kolmogorov-Smirnov (K-S) test and the Anderson-Darling (A-D) test are some of the goodness-of-fit tests used in diagnostics for stability. There are no closed-form densities for stable distributions, except for special cases such as the Normal, Cauchy and Lévy distributions.

6.5.1 Kolmogorov-Smirnov test

The K-S goodness-of-fit test makes a comparison between a fitted cdf $\hat{F}(x)$ with an empirical cdf $F_n(x)$ in order to assess the suitability of the fit. The empirical cdf $F_n(x)$ is the proportion of the observations X_1, X_2, \dots, X_n that are less than or equal to x . $F_n(x)$ is defined as

$$F_n(x) = \frac{I(x)}{n}$$

where n is the sample size and $I(x)$ is the number of X_i 's $\leq x$ (Evans et al., 2008). The K-S statistic D_n is the largest vertical distance between $\hat{F}(x)$ and $F_n(x)$ for all values of x

$$D_n = \sup_x |F_n(x) - \hat{F}(x)|. \quad (6.39)$$

The D_n statistic can be computed from

$$D_n^+ = \max_{i=1,2,\dots,n} \left\{ \frac{i}{n} - \hat{F}(X_{(i)}) \right\}, \quad D_n^- = \max_{i=1,2,\dots,n} \left\{ \hat{F}(X_{(i)}) - \frac{i-1}{n} \right\} \quad (6.40)$$

where $X_{(i)}$ is the i th order statistic and letting

$$D_n = \max\{D_n^+, D_n^-\} \quad (6.41)$$

6.5.2 Anderson-Darling test

The Anderson-Darling (A-D) test is a tail-weighted statistic where more weight is attributed to the tails and less weight in the center of the distribution. The A-D test statistic is distribution-free and is defined as

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{\left(F_n - \hat{F}(x)\right)^2}{\hat{F}(x) [1 - \hat{F}(x)]} d\hat{F}(x) \quad (6.42)$$

where n is the sample size.

The computational formula for the A-D statistic is

$$A_n^2 = - \sum_{i=1}^n \frac{2i-1}{n} \left(\ln(\hat{F}(x_{(i)})) + \ln(1 - \hat{F}(x_{(n+1-i)})) \right) - n \quad (6.43)$$

(Evans et al., 2008).

Chapter 7

Risk measures and backtesting

“Disclosure of quantitative measures of market risk, such as VaR, is enlightening only when accompanied by a thorough discussion of how the risk measures were calculated and how they related to actual performance.” Greenspan (1996)

Financial institutions set aside an amount of risk capital as per the Basel Accord and there is a direct link to the level of portfolio risk. VaR is a benchmark measure for assessing such risk. VaR aims to evaluate the maximum possible loss for a portfolio over a specified time period and its VaR calculations focus on the tails of a distribution. This provides procedures for testing the robustness of a model. Three important components in VaR are confidence level, period and loss of potential value (Christoffersen, 1998).

7.1 Current regulatory framework

The Third Basel Accord (Basel III Accord) is a regulatory framework which consists of a comprehensive set of reform measures developed for banking supervision by the Basel Committee to reinforce regulation, supervision and risk in the banking industry.

The Basel Committee is a global standard-setter for the prudential regulation of banks wherein a forum for cooperation on banking supervisory matters is provided.

It is imperative to support the regulation, practices and supervision of banks worldwide in order to enhance financial stability.

Basel III aims to:

- (i) Improve the banking sector's ability to absorb shocks arising from financial or economic stress.
- (ii) Improves risk management and governance.
- (iii) Toughen transparency and disclosures in banks.

The reform targets:

1. Bank level (micro-prudential) - Helps raise the resilience of financial institutions during periods of stress.
2. Macro-prudential: Risks that are built up across the banking sector as well as the procyclical amplification of these risks over time (BiiiCPA, 2016).

7.1.1 South African implementation of Basel III Accord

In January 2013, South Africa implemented regulations that fall in line with the Basel III framework. This aims to essentially address both bank-specific and broader systemic risks by:

- Increasing the quality of capital, with focus on common equity and the quantity of capital to ensure banks are able to efficiently absorb losses.
- Enhancing the risk coverage of the regulatory framework which includes exposures related to credit risk.
- The introduction of capital buffers which ideally should be built up in flourishing times so that in periods of stress they can be withdrawn.
- A leverage ratio should be introduced to serve as a backstop to risk-based capital requirement and this should be used to prevent the build-up of excessive leverage in the financial system.
- Raising the standards for supervision, risk management and public disclosures.

- Introducing the monitoring of proposed minimum liquidity standards to improve banks’ resilience to short-term stress and improve longer term funding.
- The introduction of additional capital buffers should be introduced for most institutions to address the issue of “too big to fail”.

The implementation period for most of the Basel III requirements were included into regulations and commenced in January 2013 and involves transitional arrangements which will be phased until 1 January 2019. Transitional arrangements are available to give banks time to meet the higher standards while still supporting lending to the economy (SARB, 2015).

7.2 Value-at-Risk

In the financial industry, the measurement of the risk of portfolio financial assets and securities are highly important. Over the years, trading practices have grown vastly with many evident situations of financial market instability leading to compromising and disastrous financial credibility. This has alarmed market participants in using reliable risk measurement techniques. One such measurement technique used is Value-at-Risk (VaR). The VaR measure is the risk measure of choice in the financial industry albeit many known imperfections. The reasoning for this choice becomes clear as one considers the associated theoretical facts and the implementation of backtesting.

Definition 7.2.1. (Value-at-Risk) (Dánielsson, 2011)

The loss on a trading portfolio such that there is a probability p of losses equaling or exceeding VaR in a given trading period and a $1 - p$ probability of losses being lower than the VaR (Dánielsson, 2011). VaR is the quantile distribution between profit and loss. The random variable \mathcal{Q} indicates the profit and loss on an investment portfolio and a particular realization is described by q . For one unit of asset, the profit and loss is indicated by

$$\mathcal{Q} = P_t - P_{t-1} \tag{7.1}$$

Generally, if the portfolio value is \mathcal{V} , then

$$\mathcal{Q} = \mathcal{V}Y, \quad (7.2)$$

where the profit/loss is the portfolio value multiplied by the returns. The density of profit/loss is denoted by $f_Q(\cdot)$. Therefore, VaR is given by

$$Pr [\mathcal{Q} \leq -\text{VaR}(p)] = p \quad (7.3)$$

or

$$p = \int_{-\infty}^{-\text{VaR}(p)} f_Q(x) dx. \quad (7.4)$$

A minus sign is used in VaR as we deal with losses.

7.2.1 Steps in VaR calculations

There are three steps in VaR calculations:

- Step 1 The probability of losses exceeding VaR, p , needs to be specified. The most common probability level is 1%. In theory, there is little guidance with regard to the choice of p but it largely depends on how the user of the risk management system prefers to interpret the VaR number.
- Step 2 This step involves the *holding period*. The holding period is defined as the time period in which losses may occur. Depending on certain circumstances this most likely occurs on Day 1. A one-day holding period is used by those who actively trade. However, longer holding period are more realistic for non-financial organizations and institutional investors. We note that the longer the holding period, the greater the value of VaR.
- Step 3 The final step is to identify the probability distribution of the profit and loss of the portfolio. This is the most complex, yet crucial facet of financial risk modeling. It is standard practice to estimate the distribution using past observations and a suitable statistical model (Dánielsson, 2011).

7.2.2 Analyzing and interpreting VaR

When interpreting VaR, it is important for one to keep in mind the probability and holding period as without these, VaR is meaningless. An identical data set may produce two different VaR estimates if different levels of VaR and holding periods are chosen. A loss suffered with probability of 1% exceeds the loss suffered by a 5% probability. VaR of a firm's portfolio is a relevant measure of risk of financial distress over a period related to the portfolio positions liquidity and the risk of extreme cash outflows. High transaction costs are caused by adverse conditions in liquidity (Danielsson, 2011).

Definition 7.2.2. (VaR violation) (Danielsson, 2011)

The VaR violation is an event such that:

$$\eta_t = \begin{cases} 1 & \text{if } y_t \leq -\text{VaR}_t \\ 0 & \text{if } y_t > -\text{VaR}_t \end{cases} \quad (7.5)$$

v_1 is the count of $\eta_t = 1$ and v_0 is the count of $\eta_t = 0$, which is obtained by:

$$\begin{aligned} v_1 &= \sum \eta_t \\ v_0 &= W_T - v_1 \end{aligned}$$

where W_T (the testing window) is the data sample over which risk is forecast; that is, the days where a VaR forecast is made.

7.3 VaR in-sample backtesting procedures

There are many shortcomings with standard VaR models such as the variance-covariance, the historical simulation and the Monte Carlo simulation. These standard VaR models are based on many assumptions. The accuracy of VaR decreases with the number of assumptions (Blanco & Oks, 2004). Hence, backtesting is a method used to address the concern of accuracy in VaR. Two methods of backtesting exist: conditional and unconditional. Conditional methods test to see if the violations are independent of each other, and conversely, unconditional methods count the number

of violations and compare them with the confidence level. A precise VaR model should satisfy the unconditional coverage and independence properties (Christoffersen, 1998). This study implements the violation ratio, Kupiec likelihood ratio test, Christoffersen conditional coverage test and the VaR duration test.

7.3.1 Violation ratio

Definition 7.3.1. (Violation Ratio) (Dánielsson, 2011)

The violation ratio is defined as:

$$VR = \frac{\text{Observed number of violations}}{\text{Expected number of violations}} = \frac{v_1}{p \times W_T} \quad (7.6)$$

Violation ratios are a quick and easy tool for checking model adequacy. If the violation ratio is greater than 1, then the VaR model *underforecasts* risk and if it is smaller than 1, the model *overforecasts* risk. The rule of thumb is that if $VR \in [0.8, 1.2]$, it is a good forecast and in the case where $VR < 0.5$ or $VR > 1.5$, the model is imprecise. This method is a relatively easy backtesting procedure to implement. However, there are some disadvantages: One which is that it cannot show the underlying causes of model failure. Therefore, we cannot solely rely on violation ratios as a mathematically justified method in determining the suitability of a model. One should use more powerful tests in checking forecast accuracy to attain formal conclusions on model adequacy, the Kupiec likelihood ratio test, Christoffersen conditional coverage test and the VaR duration test allow for formal inferences of robustness.

7.3.2 Kupiec likelihood ratio test

The Kupiec test, or otherwise referred to as the proportion of failure (POF) test, proposed by Kupiec (1995) is one of the most popular tests. The Kupiec approach tests the unconditional coverage (UC) property. We validate the accuracy of VaR model by taking note of the failure rate, that is, the proportion of times VaR is exceeded in a given sample. Denote the number of exceptions by (x) and the total number of observations by N , then the rate of failure is defined as $\frac{x}{N}$. If we suppose a

VaR number is reported at the confidence interval c , then an exception occurs if the realized loss exceeds the VaR number. Hence, the expected number of exceptions x in a total of N observations is $(1 - c)N$ (Katsenga, 2013). Undoubtedly, the number of exceptions will not always be precisely $(1 - c)N$ and it varies within an acceptable range. The range of x can be calculated in the backtesting method. Therefore, the VaR model may be accepted or rejected (Campbell, 2006).

The parameters required for VaR model backtesting using the Kupiec test is number of exceptions, x , total number of observations (N) and the confidence level (c). We define the null hypothesis H_0 as

$$H_0 : \frac{x}{N} = \frac{x^*}{N}$$

where $\frac{x}{N}$ is the expected failure rate at a given confidence level and $\frac{x^*}{N}$ is the observed failure rate. The Kupiec test is conducted just as a likelihood ratio (LR) test which is given by

$$L_{UC} = -2\ln \left[(1 - p)^{(N-x)} p^x \right] + 2\ln \left[\left(1 - \frac{x}{N} \right)^{N-x} \left(\frac{x}{N} \right)^x \right]$$

where $p = (1 - c)$. Under H_0 , the test statistic is given by the above equation and follows a chi-squared distribution with 1 degree of freedom, $(\chi^2(1))$. If the value of L_{UC} statistic falls below the critical value of $(\chi^2(1))$, then the model is adequate. Higher values above the critical region indicate a model containing inaccuracies and should, therefore, be rejected (Katsenga, 2013).

7.3.3 Christoffersen conditional coverage test

There is a major short-coming of the unconditional coverage of the Kupiec test to detect clustering of exceptions. Many tests have been suggested which examine the independence property of VaR violations in great detail.

Christoffersen (1998) suggests a conditional coverage test that examines the independence property, or exception clustering, and is also referred to as the independence test

or Markov test. Christoffersen (1998) checks if the probability of the VaR violation on any given day depends on the outcome of the previous day. Hypothetically, if the likelihood of the VaR exception increased on a preceding day, a previous VaR exception, then this indicates a need to raise the VaR level estimates since subsequent losses would imply higher risk exposure. A similar likelihood ratio statistical testing framework is applicable just as in Kupiec test for the independence of exceptions.

The test procedure is as follows:

- (i) Suppose we have data of a portfolio for N days.
- (ii) For each day, introduce a deviation indicator as below:

$$\text{Indicator}(I_t) = \begin{cases} 0 & \text{if VaR is not breached;} \\ 1 & \text{otherwise.} \end{cases}$$

Therefore, we have a sequence I_t of 0's and 1's. That is, for any two consecutive days, there are four possible outcomes $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

- (iii) Define $N_{i,j}$ ($i = 0, 1; j = 0, 1$) as the number of days in which state j occurred in one day while it was in state i the previous day. $N_{0,0}$ indicates the number of days that the previous day's indicator is 0 and the subsequent day's indicator is 0. A similar description may be used for the other possible outcomes (Jorion, 2007).

We proceed with the test by constructing a 2×2 contingency table with all possible outcomes of the deviation indicator.

	$I_{t-1} = 0$	$I_{t-1} = 1$	
$I_t = 0$	N_{00}	N_{10}	$N_{00} + N_{10}$
$I_t = 1$	N_{01}	N_{11}	$N_{01} + N_{11}$
	$N_{00} + N_{01}$	$N_{10} + N_{11}$	N

Table 7.1: Deviation Indicator outcomes (Christoffersen, 1998)

Also, define π_i as the probability of having an exception conditional on state i on the previous day. Let π_0 be the conditional probability of $(0, 1)$ occurring if the previous day is 0 and π_1 be the conditional probability of $(1, 1)$ occurring if the previous day is 1. It follows that

$$\pi_0 = \frac{N_{01}}{N_{00} + N_{01}}, \quad \pi_1 = \frac{N_{11}}{N_{10} + N_{11}}, \quad \pi = \frac{N_{01} + N_{11}}{N_{00} + N_{01} + N_{10} + N_{11}}$$

where the sum of π_0 and π_1 is π . We test H_0 : Exceptions are independent across all days ($\pi = \pi_0 = \pi_1$); that is, the probability of an exception occurring after a day of no exception is the same as occurring after a day of an exception (Campbell, 2006). If we notice that the proportions differ from each other, then this prompts one to consider the validity of the VaR measure into consideration.

If the model is accurate, then a VaR violation today should not be dependent on whether or not a violation occurred on the previous day (Jorion, 2007). The test statistic for the independence of exceptions is given as a likelihood ratio:

$$\text{LR}_{\text{Ind}} = -2\ln \left[(1 - \pi)^{(N_{00} + N_{10})} \pi^{(N_{01} + N_{11})} \right] + 2\ln \left[(1 - \pi_0)^{N_{00}} \pi_0^{N_{01}} (1 - \pi_1)^{N_{10}} \pi_1^{N_{11}} \right] \quad (7.7)$$

Just as in the Kupiec test, the LR_{Ind} -statistic follows a chi-square distribution with 1 degree of freedom ($\chi^2(1)$). In the same way, if the value of the LR_{Ind} -statistic falls below the critical value of the chi-squared distribution with 1 degree of freedom, then the specified model is correct and therefore accepted, or is otherwise rejected (Katsenga, 2013).

7.3.4 VaR duration test

The reasoning behind the duration-based tests implies that clustering of violations will result in an excessive number of relatively short and relatively long no-hit durations corresponding to market turbulence and calm periods in the market, respectively.

We define the duration of time (in days) between two VaR violations, that is, the

no-hit duration as

$$D_i = t_i - t_{i-1} \quad (7.8)$$

where t_i denotes the day of violation number i .

The duration of time between VaR violations (no-hits) should ideally be independent and not cluster. Christoffersen & Pelletier (2004) explains that under the null hypothesis that the risk model is correctly specified, the no-hit duration should have no memory and a mean duration of $1/p$ days. The no-memory property in the null hypothesis is verified by the following discrete probability distribution :

$$\begin{aligned} Pr(D = 1) &= p \\ Pr(D = 2) &= (1 - p)p \\ Pr(D = 3) &= (1 - p)^2 p \\ &\vdots \\ Pr(D = d) &= (1 - p)^{d-1} p \end{aligned}$$

A duration distribution is understood by its hazard function, which has the definition of the probability of getting a violation on day D after $D-1$ days have passed without a violation. The probability distribution shows a flat discrete hazard function as follows:

$$\begin{aligned} \lambda(d) &= \frac{Pr(D = d)}{1 - \sum_{j < d} Pr(D = j)} \\ &= \frac{(1 - p)^{d-1} p}{1 - \sum_{j=0}^{d-2} (1 - p)^j p} \\ &= p \end{aligned}$$

The memory-free continuous random distribution is shown as the exponential, that is, under the null hypothesis, the no-hit durations should be represented as

$$f_{\text{exp}}(D; p) = p e^{-pD} \quad (7.9)$$

Consider the Weibull distribution where

$$f_W(D; a, b) = a^b b D^{b-1} \exp(-(aD)^b) \quad (7.10)$$

The advantage of the Weibull distribution is that the hazard function is in closed-form

$$\lambda_W(D) = \frac{f_w(D)}{1 - F_W(D)} = a^b b D^{b-1} \quad (7.11)$$

where the exponential distribution occurs as a special case with a flat hazard, when $b = 1$. When $b < 1$, the Weibull has a decreasing hazard function. This corresponds to a large number of very short durations which implies very volatile periods as well as a large number of long durations (periods of tranquility). There may be evidence of misspecified volatility dynamics in the model.

VaR violation clustering in the null hypothesis is of great interest. Therefore, we explicitly test the null hypothesis

$$H_{0,ind} : b = 1 \quad (7.12)$$

The Gamma distribution can be used under the alternative hypothesis. The density is given by

$$f_r(D; a, b) = \frac{a^b D^{b-1} \exp^{-aD}}{\Gamma(b)} \quad (7.13)$$

which nests the exponential when $b = 1$. In this case, we have the independence test in the null hypothesis as

$$H_{0,ind} : b = 1 \quad (7.14)$$

The Gamma distribution does not have a closed-form solution for the hazard function, but the first and second moments are $\frac{b}{a}$ and $\frac{b}{a^2}$, respectively. Excess dispersion which is defined as the variance over the squared expected value is $\frac{1}{b}$ in the case of the Gamma distribution. The average duration in the exponential distribution is $\frac{1}{p}$ and the variance is $\frac{1}{p^2}$. Thus, excess dispersion is 1 in the exponential case (Christoffersen & Pelletier, 2004).

Chapter 8

Analysis of BRICS financial data

This chapter explores an empirical analysis of the returns of BRICS countries to U.S. Dollar exchange rates. R is the statistical software that was used with the following packages: fbasics, TSA, Stats, time series, ADGofTest, rugarch, VarES, QRM and STABLE. The STABLE package was made available from Prof. John Nolan's website: www.RobustAnalysis.com.

8.1 Exchange rate data

The exchange rate data for each of the BRICS countries were obtained from the Board of Governors of the Federal Reserve System (Central Bank of the United States) and the Bank of Russia. This study covers the time period of January 2011 to January 2016 and the currency of each country is the Brazilian Real (BRL), Russian Ruble (RUB), Indian Rupee (INR), Chinese Yuan Renminbi (CHY) and South African Rand (ZAR) will be compared against the U.S. Dollar.

8.2 Descriptive Statistics

8.2.1 Time series plots of daily exchange rates

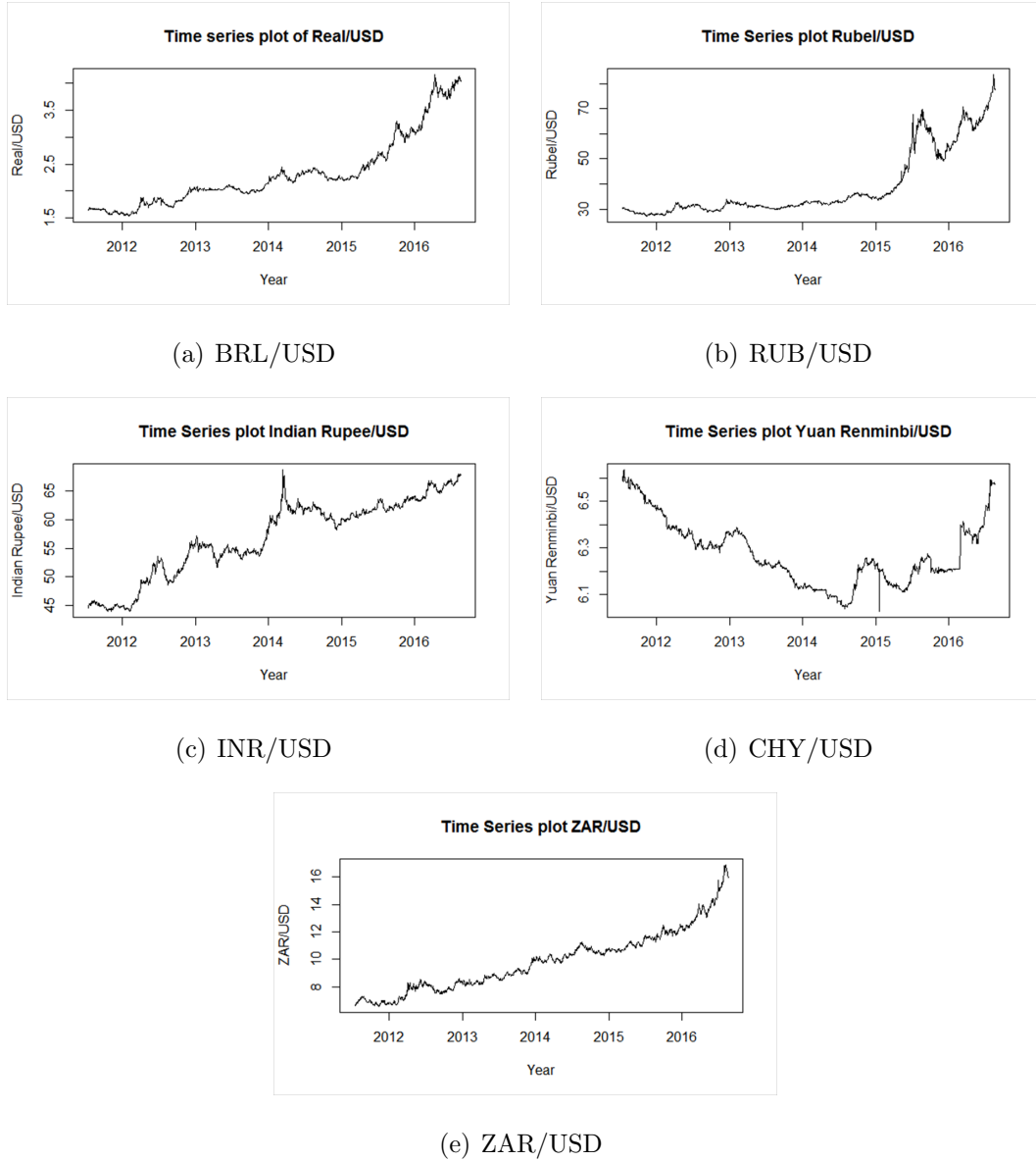


Figure 8.1: Time series plots of BRICS to USD daily exchange rates

The time series plots in Figure 8.1 indicate non-stationarity and a general upward trend in the returns of exchange rates except for the CHY/USD where a downward trend is visible followed by an upward trend starting in the latter part of 2014. Heteroscedasticity is also visible.

8.3 Calculating log-returns

To obtain the exchange rate returns, the natural logarithm of exchange rates are differenced, i.e

$$Y_t = \log(P_t) - \log(P_{t-1})$$

where Y_t is the log-return on day t , $\log(P_t)$ is the natural logarithm of the present day exchange rate and $\log(P_{t-1})$ is the natural logarithm of exchange rates on the previous day.

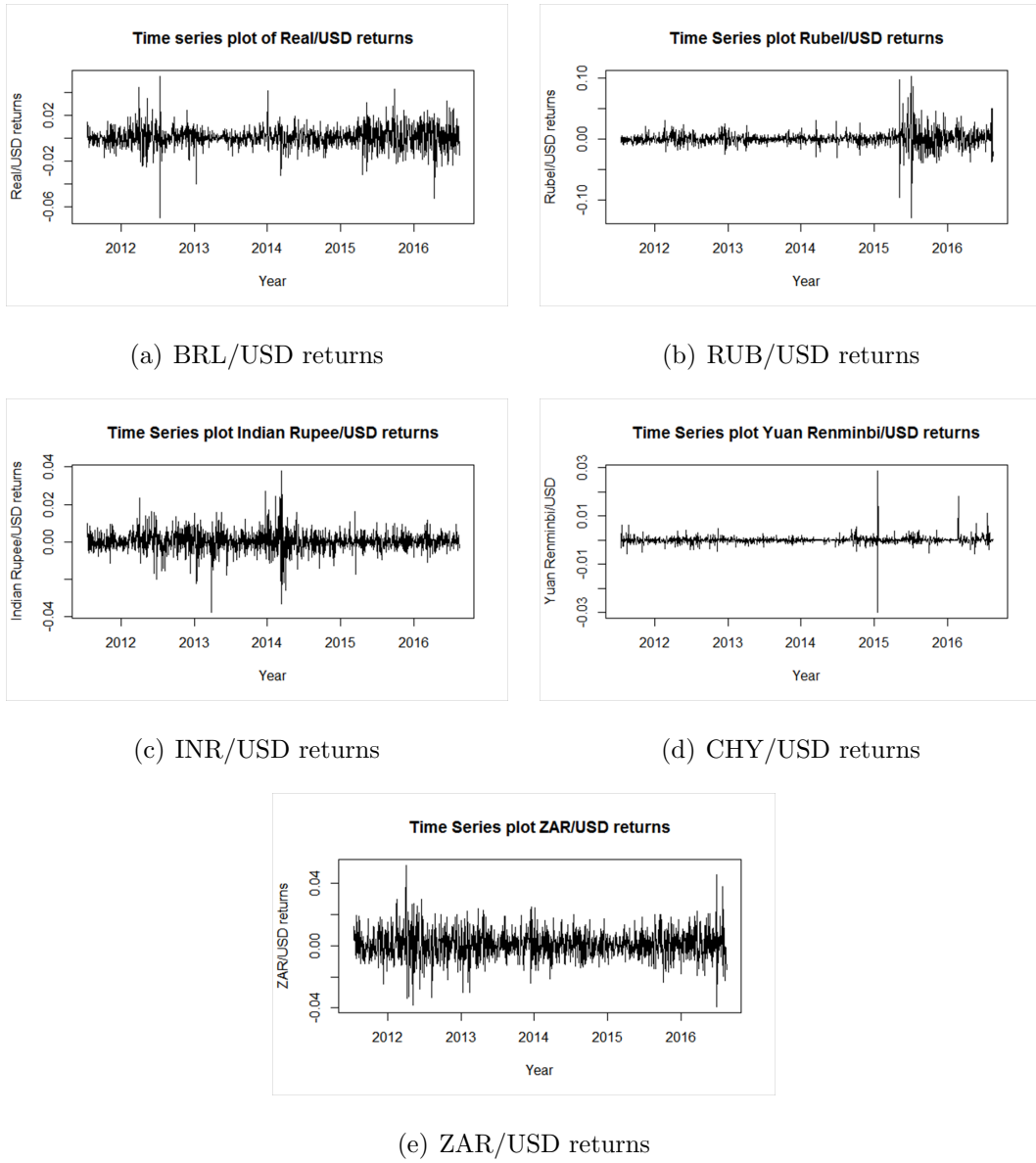


Figure 8.2: Time series plots of BRICS to USD exchange rate returns

Figure 8.2 shows the time series plots of the log-returns. From Figure 8.2, we see that the plots now indicate that all the returns are stationary and it seems volatility clustering is visible. However, we notice a unique case in the CHY where the presence of outliers is also visible.

Table 8.1: Descriptive summary statistics of daily return of BRICS to USD exchange rates.

	Brazil	Russia	India	China	South Africa
No. of obs.	1324	1324	1324	1324	1324
Minimum	-0.069580	-0.128638	-0.037560	-0.029921	-0.039134
Maximum	0.053943	0.102993	0.037919	0.028648	0.051464
1-Quartile	-0.004055	-0.004105	-0.002345	-0.000538	-0.005011
3-Quartile	0.005683	0.004887	0.003029	0.000444	0.005752
Mean	0.000678	0.000702	0.000317	-0.000002	0.000665
Median	0.000432	0.000292	0.000152	0	0.000146
Sum	0.897491	0.929731	0.420084	-0.002172	0.880943
SE Mean	0.000254	0.000348	0.000154	0.000048	0.000252
LCL Mean	0.000180	0.000019	0.000015	-0.000096	0.000171
UCL Mean	0.001176	0.001385	0.000620	0.000093	0.001160
Variance	0.000085	0.000161	0.000031	0.000003	0.000084
Stdev	0.009238	0.012673	0.005611	0.001752	0.009174
Skewness	-0.220277	0.060056	0.016852	0.779042	0.236834
Excess Kurtosis	5.533864	20.088195	6.591997	127.498816	2.217928
Jarque-Bera statitic	1708.03	22339.96	2407.86	899699.1	285.72
(<i>p</i>-value)	$< 2.2e^{-16}$	$< 2.2e^{-16}$	$< 2.2e^{-16}$	$< 2.2e^{-16}$	$< 2.2e^{-16}$

Table 8.1 summarizes the descriptive statistics of the daily returns of BRICS to USD exchange rates. We notice positive mean values for Brazil, Russia, India and South Africa, which implies exchange rates were increasing very slightly over the time

period of 2011-2016. Conversely, the negative mean for China's currency implies that the exchange rates exhibits a slight decreasing trend. Excess kurtosis is evident for all countries which suggests a leptokurtic distribution which is sharper with values concentrated around the mean with longer, thicker tails. The Jarque-Bera test rejects the hypothesis of Normality at all levels of significance.

8.4 Tests for autocorrelation

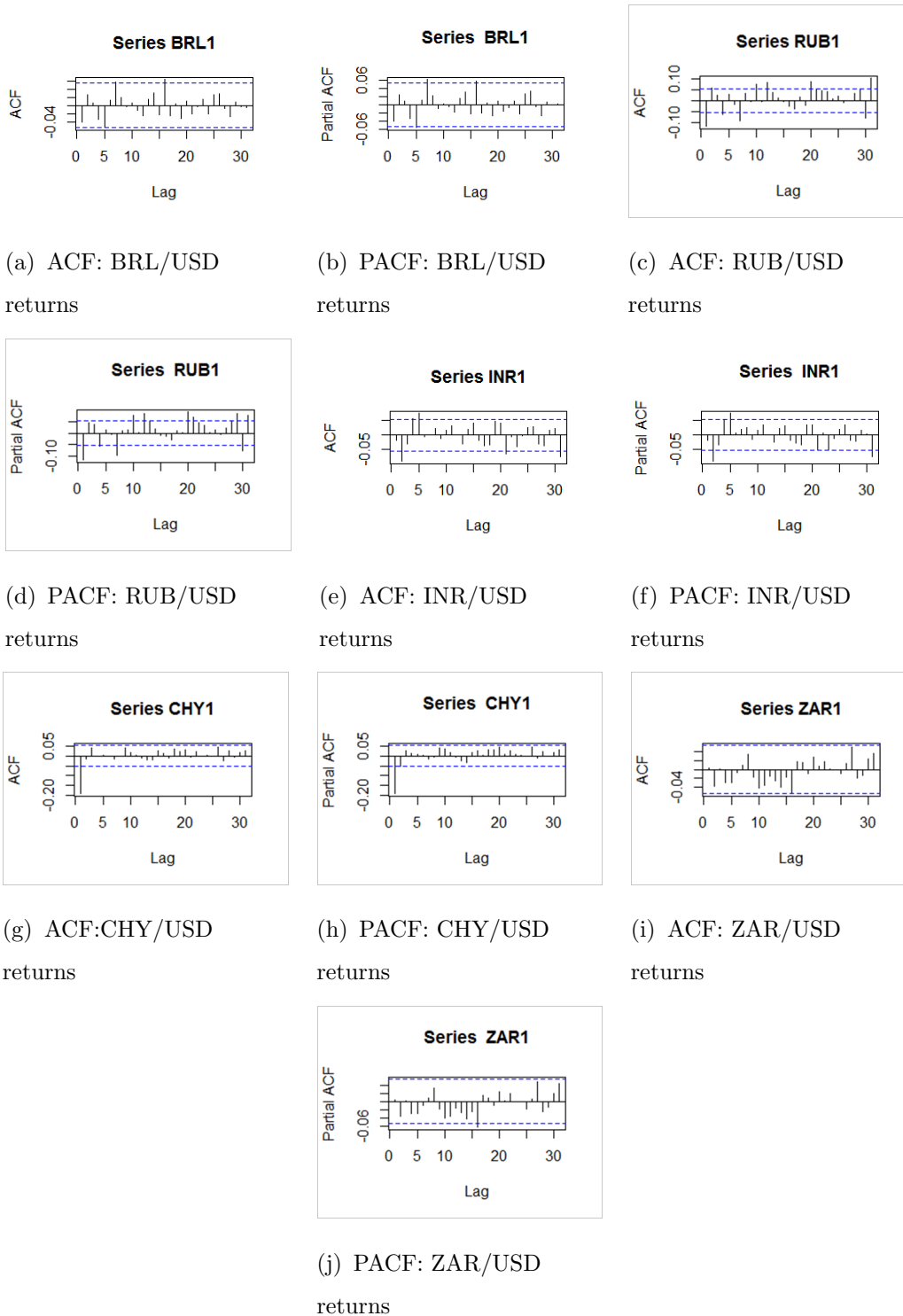


Figure 8.3: ACF and PACF plots of daily BRICS to USD exchange rate returns

The ACF and PACF plots in Figure 8.3 show that the daily returns exhibit significant serial correlation for all the BRICS countries. For Brazil, Russia, India, China and South Africa: the ACF plots have significant serial correlation at lags 5, 1, 2, 1 respectively. Likewise, the PACF plots have significant serial correlation at lags 5, 1, 2, 1 and 16.

8.4.1 Ljung-Box test for autocorrelation

We formally test for the presence of serial correlation in the returns. Table 8.2 shows the results of the Ljung-Box test at lag 10.

Table 8.2: Ljung-Box test of daily returns for BRICS to USD exchange rates

	Brazil	Russia	India	China	South Africa
Test Statistic	13.466	49.8014	25.9912	55.5038	8.6774
p -value	0.1988	$2.903e^{-07}$	0.003752	$2.542e^{-08}$	0.563

In Table 8.2 we notice at 5% level of significance we favor the null hypothesis for Brazil and South Africa implying that the returns for the Real and Rand to the USD exchange rate are independent. Conversely, we reject the null hypothesis for Russia, India and China implying that the Ruble, Rupee and Yuan Reminbi to USD exchange rate exhibit serial correlation. This test provides mixed results on the hypothesis of independence in the foreign exchange rate market. The fitting of a statistical distribution usually assumes that the returns are i.i.d, that is, there is no serial correlation and no heteroskedasticity. However, the empirical properties of financial returns as noted by McNeil et al. (2005) describes that some returns in financial data show serial correlation.

8.5 Tests for stationarity

In order to fit GARCH models, the returns of the exchange rates need to be stationary. We test for stationarity using the ADF, P-P and KPSS tests. Table 8.3 shows the results for testing stationarity.

Table 8.3: Results for ADF, PP and KPSS unit root tests for BRICS/USD Exchange rate returns

		Brazil	Russia	India	China	South Africa
ADF	Test Statistic	-10.7538	-10.7919	-9.9134	-10.6831	-11.9453
	p -value	0.01	0.01	0.01	< 0.01	< 0.01
PP	Test Statistic	-1364.944	-1494.782	-1304.498	-1536.263	-1246.158
	p -value	0.01	0.01	0.01	< 0.01	< 0.01
KPSS	Test Statistic	0.2543	0.5316	0.069	0.9898	0.08
	p -value	0.1	0.03454	0.1	< 0.01	0.1

Table 8.3 shows the test statistics and corresponding p -values for the ADF, PP and KPSS stationary tests. For the ADF and PP tests, stationarity is defined in the alternative hypothesis. We reject the null hypothesis in favor of the alternative hypothesis since the p -values are less than 0.05 and conclude that the returns of daily exchange rates exhibit stationarity. For the KPSS test, stationarity is defined in the null hypothesis. We notice that the exchange rates for Brazil, India and South Africa exhibit stationarity as the observed p -values = 0.1 are greater than 0.05, and is indicative of favoring the null hypothesis.

8.5.1 Checking for hidden periodicity

In economic time series data, we may frequently assume that data contains cyclical phenomena. Therefore, in this section, the SAS procedure PROC SPECTRA is used to check for hidden periodicities in exchange rate data. The Fisher's Kappa test checks for the presence of a periodic component. The null hypothesis suggests that the series is a white noise and the alternative hypothesis suggests that the series contains a periodic component at the largest periodogram ordinate.

Table 8.4: Fisher's Kappa test for detecting hidden periodicity in exchange rates.

	Brazil	Russia	India	China	South Africa
m	662	662	662	662	662
Fisher's Kappa test statistic	260.2005	247.9837	353.8839	504.1051	323.3499
Critical value: $C_m(0.05)$	9.313	8.889	9.313	9.313	9.313
Decision	Contains a periodic component	Contains a periodic component	Contains a periodic component	Contains a periodic component	Contains a periodic component

Table 8.4 shows that the Fisher's Kappa test statistic $> C_m(0.05)$ for all BRICS countries. We reject the null hypothesis and conclude that the series contains a periodic component at the largest periodogram ordinate. The period for each country was found to be 1325 which equals the number of observations. It is recommended that the results require further in-depth analysis (which will not be carried out in this study) as this phenomenon is yet to have an explanation.

8.6 Covariance matrix

$$\mathbf{\Gamma}_k = \begin{bmatrix} 8.53e^{-05} & -1.34e^{-06} & 1.38e^{-05} & 9.32e^{-07} & 4.23e^{-05} \\ -1.34e^{-06} & 0.0002 & -1.63e^{-06} & -1.32e^{-07} & 3.37e^{-08} \\ 1.38e^{-05} & -1.63e^{-06} & 3.15e^{-05} & 1.03e^{-06} & 2.04e^{-05} \\ 9.32e^{-06} & -1.32e^{-07} & 1.03e^{-06} & 0.0006 & 1.28e^{-06} \\ 4.23e^{-05} & 3.37e^{-08} & 2.04e^{-05} & 1.28e^{-06} & 8.42e^{-05} \end{bmatrix}$$

A positive covariance indicates that both variables either increase or decrease together and a negative covariance implies that if one increases, the other decreases (or vice-versa) with regard to the rate of exchange.

8.7 Correlation matrix

$$\rho_k = \begin{bmatrix} 1 & -0.012 & 0.266 & 0.06 & 0.499 \\ -0.012 & 1 & -0.023 & 0.01 & 0.0002 \\ 0.266 & -0.023 & 1 & 0.1 & 0.397 \\ 0.06 & 0.01 & 0.1 & 1 & 0.08 \\ 0.499 & 0.0002 & 0.397 & 0.08 & 1 \end{bmatrix}$$

The cross-correlation matrix shown above generally indicates a slight positive correlation amongst the BRICS countries exchange rates.

8.8 Stable parameter estimation

In this study, stable parameters are estimated under Nolan's S_0 -parameterization using the maximum likelihood estimation method. Table 8.5 summarizes the values estimated for each stable parameter for countries in BRICS.

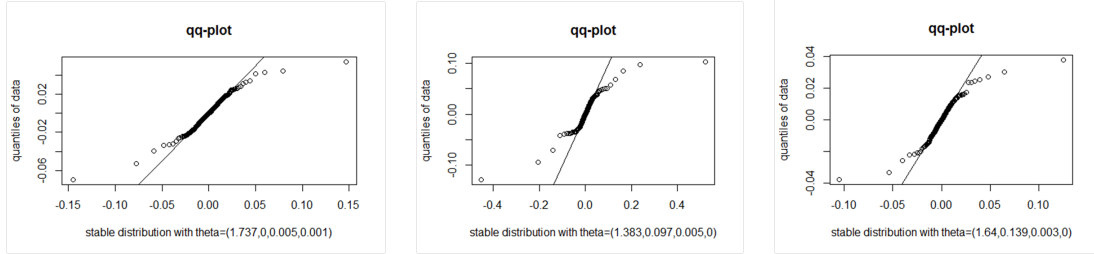
We proceed to fit Nolan's $S_0(\alpha, \beta, \gamma, \delta)$ univariate stable distributions to the daily returns of each countries exchange rates to the USD using the parameter estimates.

Table 8.5: Stable parameter estimates for daily returns under Nolan's S_0 -parameterization.

	Brazil	Russia	India	China	South Africa
α	$1.736639e^{+00}$	1.3829602460	1.6396894583	$1.386544e^{+00}$	1.7997024653
β	$-2.102931e^{-08}$	0.0970088632	0.1387799128	$1.752244e^{-09}$	0.3483880044
γ	$5.323732e^{-03}$	0.0046774792	0.0029549139	$5.161903e^{-04}$	0.0057011342
δ	$6.868068e^{-04}$	0.0002193559	0.0001643851	$-4.192556e^{-05}$	0.0002034499

8.9 Goodness of fit tests and diagnostics

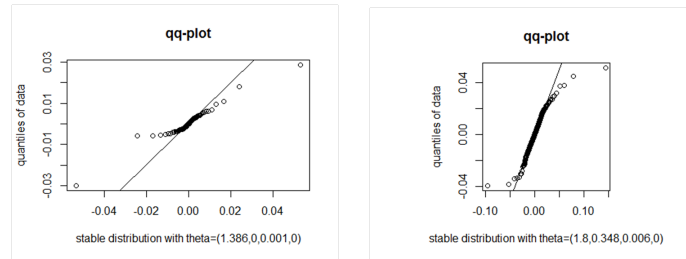
8.9.1 Q-Q plots of BRICS exchange rate returns



(a) BRL/USD returns

(b) RUB/USD returns

(c) INR/USD returns



(d) CHY/USD returns

(e) ZAR/USD returns

Figure 8.4: Q-Q plots of BRICS to USD returns

Figure 8.4 shows the Q-Q plot and variance stabilized plots of BRICS to USD exchange rates. The Q-Q plots in the figure appear to be visually compressed with extreme values dominating the plot. The heavy tails, evident in these Q-Q plots, indicate that the extreme order statistics have a lot of variability. Therefore, deviations from the ideally straight line Q-Q plot are difficult to assess. This also shows that extreme tails from the data set are lighter than the stable model (Nolan, 2005). We notice that the Q-Q plots imply the inadequacy of the stable model at extreme values. Therefore, the problems mentioned about the Q-Q plots lead us to focus on the variance stabilized P-P plots.

8.9.2 Variance stabilized P-P plots of BRICS exchange rate returns

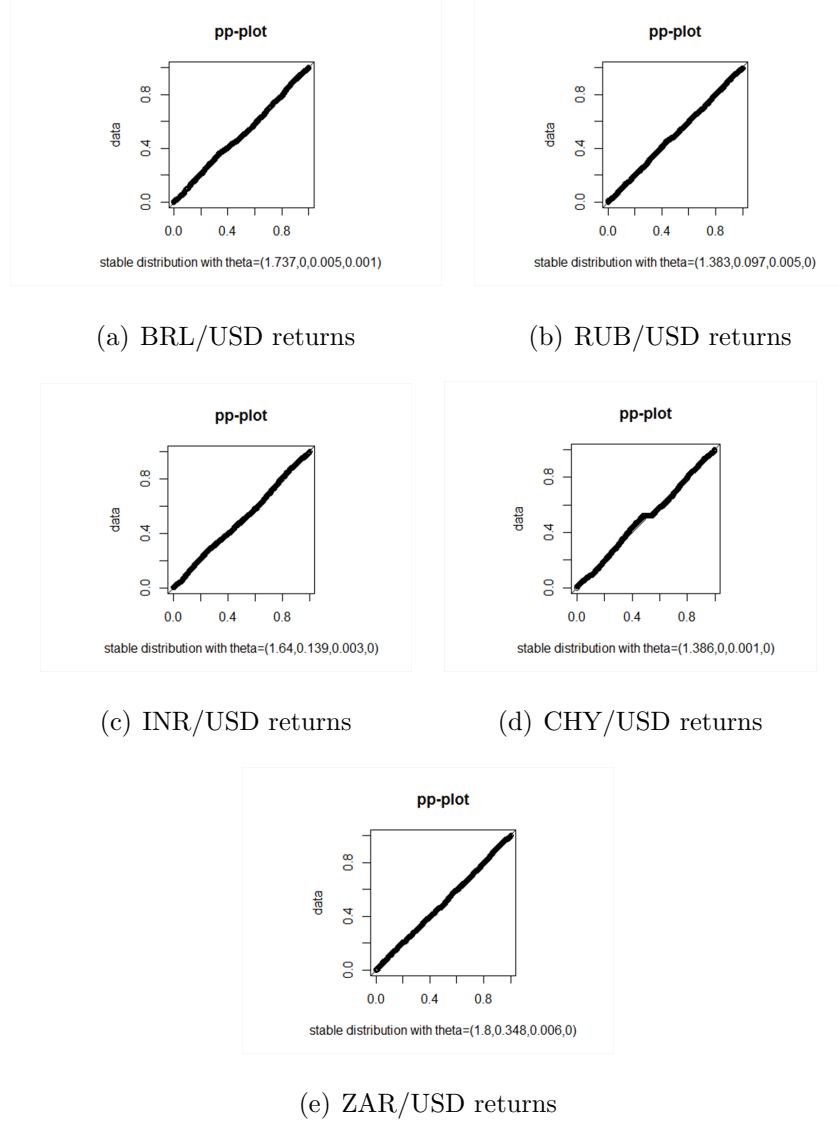


Figure 8.5: Q-Q plots of BRICS to USD returns

The variance stabilized P-P plots allow for comparisons over the entire range of the data. One may notice a horizontal line segment in the P-P plot of sub-figure (d) in Figure 8.5, this results from days where the exchange rate remained unchanged on successive days. The figures show the adequacy of the fitted univariate models for all BRICS countries excluding China where slight deviations evident in the plots suggest that the fitted univariate stable model may be adequate for modeling the

returns of the Chinese Yuan Renminbi, however a possible alternative heavy tailed model may be more robust for modelling the Chinese exchange rates.

We use the K-S and A-D tests to check for model adequacy.

Table 8.6: Goodness-of-fit tests of daily returns

		Brazil	Russia	India	China	South Africa
K-S	Test Statistic	0.02640462	0.03058174	0.02716278	0.04397026	0.01893238
	p -value	0.3105694	0.1651808	0.2789941	0.01154931	0.7258011
A-D	Test Statistic	1.0244	0.3168	1.0234	2.0882	0.4064
	p -value	0.3447	0.925	0.3452	0.08219	0.8423

We define the null hypothesis to be:

Daily returns of BRICS to USD exchange rates follow a univariate $S_0(\alpha, \beta, \gamma, \delta)$ model

Table 8.6 indicates that at a 5% level of significance, the p -values for the Kolmogorov-Smirnov test indicate that we favor the null hypothesis for all countries in BRICS except for China where the p -value = 0.01154931 < 0.05. However, at a 10 % level of significance, for China, where the p -value = 0.08219 favors the null hypothesis. We notice from the Anderson-Darling test that all fitted stable models are significant. From Table 8.6, we conclude that the fitted univariate $S_0(\alpha, \beta, \gamma, \delta)$ model for each country is adequate in describing the daily exchange rate returns for each country.

8.9.3 Univariate stable density plots

Empirically, we compare the densities of BRICS exchange rate returns to univariate S_0 stable distribution. Figure 8.6 shows graphically a close fit of the estimated univariate S_0 model to the daily returns of BRICS exchange rates as the fitted stable model does not deviate much from the returns of the exchange rates. A better fit for the data is provided over most of the range with extreme tails being overestimated.

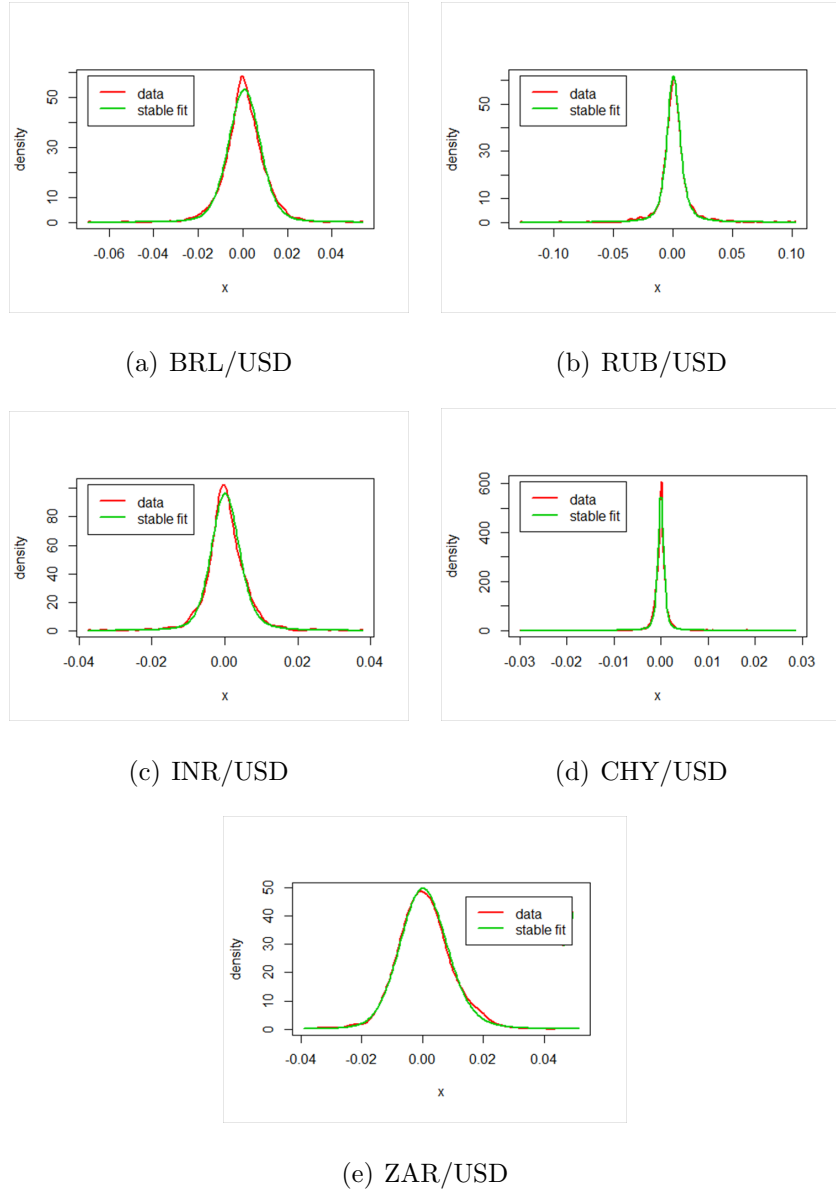


Figure 8.6: Stable density plots of BRICS to USD

8.10 VaR estimates and backtesting

We estimate VaR at 1%, 5%, 95% and 99% levels.

VaR Estimates				
	1%	5%	95%	99%
Brazil	-0.02487615	-0.01307162	0.01444523	0.02624979
Russia	-0.04346203	-0.01496782	0.01716632	0.05048657
India	-0.01556429	-0.007493927	0.008698845	0.01868417
China	-0.005188648	-0.001807187	0.001724221	0.005105673
South Africa	-0.02103707	-0.01309692	0.01552886	0.02783011

Table 8.7 presents VaR estimates for each BRICS country using a fitted stable model.

	Violation Ratio				<i>p</i> -value of Kupiec Test				<i>p</i> -value of Christoffersen's Test			
	1%	5%	95%	99%	1%	5%	95%	99%	1%	5%	95%	99%
Brazil	0.76	1.04	0.1	1.06	0.35	0.73	0.55	0.21	0.34	0.09	0.45	0.21
Russia	0.23	1.07	0.1	1.01	0.001	0.55	0.63	0.01	0.001	0.001	0.03	0.008
India	0.98	1.21	0.1	1.00	0.95	0.09	0.47	0.06	0.48	0.02	0.12	0.01
China	0.03	0.92	0.99	1.00	0.003	0.51	0.28	0.35	0.05	0.01	0.05	0.004
South Africa	1.36	0.94	0.99	1.01	0.21	0.59	0.28	0.06	0.24	0.74	0.52	0.014

The interpretation of violation ratios implies that it is a good forecast and the stable model is a good fit. We observe at a 5% level of significance, the Kupiec test indicates that the fitted univariate stable model is a good fit at higher VaR levels for all the mentioned BRICS countries. Also, we notice high *p*-values on all VaR levels for the Kupiec test for the Brazilian Real and the South African Rand daily returns. This indicates the suitability of the stable model for these exchange rate data sets. The fitted stable models for Russia, India and China may also be considered as fairly good. The Christoffersen test shows high *p*-values at 5% and 95% for Brazil and South Africa. For Christoffersen test, at almost all VaR levels for Russia, India and China, the clustering of VaR violations occur as indicated by the observed lower *p*-values. This is caused by the number of observations that exceed VaR estimates.

Table 8.7: VaR duration test

	VaR Duration Test			
	1%		5%	
	Test Statistic	Decision	Test Statistic	Decision
Brazil	0.13	Fail to Reject H_0	0.001	Reject H_0
Russia	0.01	Reject H_0	$2.81e^{-0.8}$	Reject H_0
India	0.004	Reject H_0	0.02	Reject H_0
China	0.47	Fail to Reject H_0	0.12	Fail to Reject H_0
South Africa	0.05	Fail to Reject H_0	0.05	Fail to Reject H_0

For China and South Africa, the null hypothesis of a correctly specified model is favored. Further analysis is required in this regard since the table shows mixed results.

8.11 Multivariate stable data analysis

We explore the possibility of fitting a multivariate stable model for BRICS to USD exchange rates.

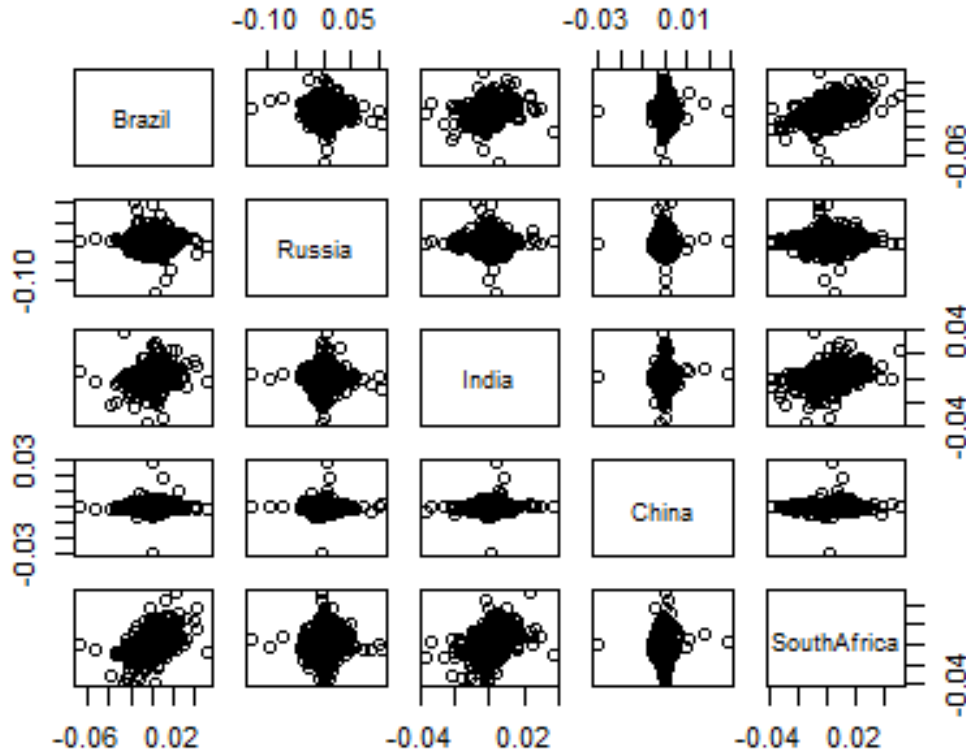


Figure 8.7: Pairwise scatterplots of BRICS countries to USD exchange rate

The pairwise scatterplots in Figure 8.7 show an elliptical pattern and in some cases, the estimated index of stability (α) is similar (From Table 8.5: Brazil and South Africa are similar as well as Russia and China) This may suggest the estimation of a jointly stable bivariate elliptical model.

8.11.1 Multivariate stable parameter estimation

We proceed by fitting bivariate elliptical stable distributions using the Rachec-Xin-Cheng method. Estimates for α , δ and R are given in Table 8.8.

α	δ	R
Brazil - Russia		
1.523953	(0.0002700612 0.0005117953)	$\begin{pmatrix} 2.422102e^{-05} & 2.301749e^{-06} \\ 2.301749e^{-06} & 2.482960e^{-05} \end{pmatrix}$
Brazil - India		
1.578767	(0.0002363954 0.0003955759)	$\begin{pmatrix} 1.487349e^{-05} & 1.487349e^{-05} \\ -5.713798e^{-07} & 1.568826e^{-05} \end{pmatrix}$
Brazil - China		
0.8267018	$(-2.761757e^{-05} \quad -1.840455e^{-05})$	$\begin{pmatrix} 1.443059e^{-06} & -6.675295e^{-08} \\ -6.675295e^{-08} & 1.546700e^{-06} \end{pmatrix}$
Brazil - South Africa		
1.777212	(0.0002797684 0.0005635676)	$\begin{pmatrix} 3.136096e^{-05} & 3.366216e^{-06} \\ 3.366216e^{-06} & 2.999531e^{-05} \end{pmatrix}$
Russia - India		
1.423015	(0.0001622021 0.0001696470)	$\begin{pmatrix} 1.302361e^{-05} & 9.005531e^{-07} \\ 9.005531e^{-07} & 1.318703e^{-05} \end{pmatrix}$
Russia - China		
0.7975918	$(1.289010e^{-05} \quad -3.398058e^{-05})$	$\begin{pmatrix} 1.368924e^{-06} & 5.427067e^{-08} \\ 5.427067e^{-08} & 1.375297e^{-06} \end{pmatrix}$
Russia - South Africa		
0.8010986	$(-4.172985e^{-06} \quad -1.675845e^{-05})$	$\begin{pmatrix} 1.362003e^{-06} & 4.625333e^{-08} \\ 4.625333e^{-08} & 1.376149e^{-06} \end{pmatrix}$
India - China		
1.037113	$(-1.857093e^{-05} \quad -7.387177e^{-05})$	$\begin{pmatrix} 1.191080e^{-06} & -7.300018e^{-08} \\ -7.300018e^{-08} & 1.157222e^{-06} \end{pmatrix}$
India - South Africa		
1.650807	$(1.056904e^{-05} \quad 1.794698e^{-04})$	$\begin{pmatrix} 1.780588e^{-05} & 1.934288e^{-06} \\ 1.934288e^{-06} & 1.745208e^{-05} \end{pmatrix}$
China - South Africa		
0.8094735	$(-4.124276e^{-05} \quad -2.924229e^{-05})$	$\begin{pmatrix} 1.620507e^{-06} & 2.634408e^{-08} \\ 2.634408e^{-08} & 1.752355e^{-06} \end{pmatrix}$

Table 8.8: Bivariate elliptical stable parameter estimation for daily returns

The returns of each BRICS country were fitted with an elliptical stable model with index of stability α , shift matrix δ and shape matrix R . An elliptical multivariate stable model allows for capturing heavy tails and dependence. Elliptical stable models also allowed accumulated returns to have the same distribution as daily returns since the cumulative sum of stable terms are always stable. The added advantage of the bivariate stable model is that linear combinations are naturally univariate stable. When R is a multiple of

the identity matrix, then the isotropic or radially symmetric case arises. From Table 8.8 there is strong empirical evidence for the isotropic heavy tailed stable model for the Brazil-South Africa model and the Russia-China model.

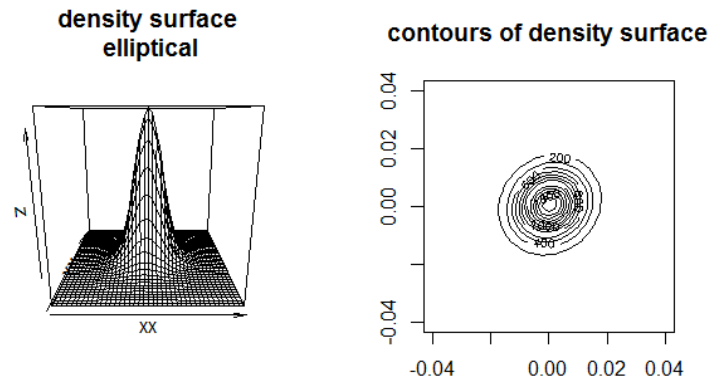


Figure 8.8: Estimated density surface and countour plot for a bivariate elliptical stable fit of Brazillian Real and South African Rand exchange rates.

Figure 8.8 shows the bivariate elliptical stable density plot of BRL-ZAR/USD exchange rate returns, the other bivariate plots for intra-BRICS countries combination are found in Appendix B. The fitted spectral measure was used to plot the fitted bivariate density shown in Figure 8.7. The spread of the spectral measure is spiky and masks a pattern that is more evident in the density surface. These are approximately the elliptical countours of the fitted density.

Chapter 9

Stable-GARCH(1,1) model

In this chapter, we investigate the modeling of BRICS exchange rate returns with stable distributions together with a GARCH(1,1) filter.

ARCH LM test for heteroscedasticity

Table 9.1: ARCH LM test for heteroscedasticity

	Brazil	Russia	India	China	South Africa
χ^2 Test Statistic	100.7833	357.9457	203.7792	435.9872	106.4271
p -value	$4.441e^{-16}$	$2.2e^{-16}$	$2.2e^{-16}$	$2.2e^{-16}$	$2.2e^{-16}$

The ARCH LM test was used to test for heteroscedasticity on the returns of the exchange rates. Table 9.1 shows the p -values based on the ARCH LM test where we see that the test confirms a strong ARCH effect in the daily BRICS exchange rate returns.

We fit a GARCH(1,1) model to capture the time-varying volatility in the exchange rate returns.

The following steps are used for calculating VaR and backtesting for the stable-GARCH(1,1) model:

- Step 1. A GARCH(1,1) model is fitted to the returns using pseudo ML procedure based on Normal assumptions.
- Step 2. Standardized residuals are extracted from the GARCH(1,1) model.
- Step 3. Univariate stable distribution is fitted to the standardized residuals using ML estimation.
- Step 4. VaR estimates are calculated and then we apply the Kupiec and Christoffersen backtesting procedures.

9.1 Stable-GARCH(1,1) model

9.1.1 Fitting a univariate GARCH(1,1) model to returns

Now, we apply a GARCH filter to the daily returns of BRICS/USD exchange rates. Table 9.2 presents the parameter estimates for the GARCH(1,1) model.

Table 9.2: GARCH(1,1) parameter estimation for daily returns

Returns										
	Brazil		Russia		India		China		South Africa	
	Estimate	p-value	Estimate	p-value	Estimate	p-value	Estimate	p-value	Estimate	p-value
α_0	$6.286e^{-07}$	0.00938	$3.500e^{-06}$	$9.30e^{-05}$	$3.806e^{-07}$	0.0163	$6.447e^{-07}$	$< 2e - 16$	$6.817e^{-07}$	0.0734
α_1	$7.453e^{-02}$	$1.62e^{-10}$	$1.977e^{-01}$	$1.91e^{-08}$	$5.072e^{-02}$	$6.76e^{-08}$	$3.388e^{-01}$	2.39e-09	$4.637e^{-02}$	$3.82e^{-07}$
β_1	$9.244e^{-01}$	$< 2e - 16$	$8.010e^{-01}$	$< 2e - 16$	$9.370e^{-01}$	$< 2e - 16$	$5.474e^{-01}$	$< 2e^{-16}$	$9.469e^{-01}$	$< 2e^{-16}$

Table 9.2 records the MLE parameter estimates for the GARCH(1,1) model with Gaussian innovations. All parameters are statistically significant.

The volatility equation can be written as:

Brazil: $\sigma_t^2 = 6.286e^{-07} + 7.453e^{-02} a_{t-1}^2 + 9.244e^{-01} \sigma_{t-1}^2$

Russia: $\sigma_t^2 = 3.500e^{-06} + 1.977e^{-01} a_{t-1}^2 + 8.010e^{-01} \sigma_{t-1}^2$

India: $\sigma_t^2 = 3.806e^{-07} + 5.072e^{-02} a_{t-1}^2 + 9.370e^{-01} \sigma_{t-1}^2$

China: $\sigma_t^2 = 6.447e^{-07} + 3.388e^{-01} a_{t-1}^2 + 5.474e^{-01} \sigma_{t-1}^2$

South Africa: $\sigma_t^2 = 6.817e^{-07} + 4.637e^{-02} a_{t-1}^2 + 9.469e^{-01} \sigma_{t-1}^2$

We assess the model adequacy for the fitted GARCH(1,1) model.

Table 9.3: Ljung-Box test for daily returns of the fitted GARCH(1,1) model

Exchange rate returns					
	Brazil	Russia	India	China	South Africa
Test Statistics					
Q(10)	4.47726	9.011007	7.24296	8.263875	7.805792
Q(15)	5.419485	21.31796	8.865122	10.01173	15.74428
Q(20)	10.40516	25.47511	11.01203	17.83088	21.49418
Q²(10)	3.979737	1.357615	4.577694	0.1271452	15.04757
Q²(15)	4.635103	2.295096	6.399088	0.1860767	22.36748
Q²(20)	6.479483	2.572245	9.301564	0.2119258	23.19972
<i>p</i> -values					
Q(10)	0.92326	0.5310593	0.7023262	0.6030792	0.6478003
Q(15)	0.9879394	0.1269741	0.8844635	0.8190012	0.3992475
Q(20)	0.9602165	0.1838551	0.9459101	0.5985479	0.3685761
Q²(10)	0.9482565	0.9993146	0.9175477	1	0.130337
Q²(15)	0.994791	0.9999268	0.9722365	1	0.09852404
Q²(20)	0.9980753	0.9999989	0.9791439	1	0.2790947

The Ljung-Box test statistics of the standardized residuals for $Q(10)$, $Q(15)$, $Q(20)$ are given in Table 9.3. The corresponding p -values are all greater than 0.05. In addition, the Ljung-Box test statistics of the squared residuals also have p -values greater than 0.05. We conclude that the GARCH(1,1) model is adequate in describing the exchange rate returns.

The ARCH LM test was used to check for ARCH effects in the standardized residuals of the GARCH(1,1) model. Table 8.4 below provides the test statistics and p -values.

	Brazil	Russia	India	China	South Africa
χ^2 Test Statistic	4.489646	1.377172	5.021196	0.1767796	16.89186
p -value	0.972896	0.9999176	0.9572673	1	0.1537105

Table 9.4: ARCH LM test for heteroscedasticity

The ARCH LM test results indicated in the above table suggest that there are no ARCH effects in the standardized residuals.

We extract the standardized residuals from the GARCH(1,1) model to conduct the analysis that follows.

Table 9.5: Descriptive statistics of residuals to the fitted GARCH(1,1) model for daily log returns

	Brazil	Russia	India	China	South Africa
No. of obs.	1324	1324	1324	1324	1324
Minimum	-7.684248	-5.362317	-8.512789	-19.273947	-3.474804
Maximum	7.049548	13.121522	5.506422	15.161924	5.211944
1-Quartile	-0.479421	-0.496931	-0.472824	-0.374174	-0.567512
3-Quartile	0.655071	0.587386	0.606421	0.293920	0.677844
Mean	0.079771	0.078142	0.069901	-0.030126	0.083222
Median	0.059205	0.034792	0.017351	0	0.017955
Sum	105.617374	103.460437	92.549547	-39.886652	110.186574
SE Mean	0.027315	0.027359	0.027397	0.027491	0.027385
LCL Mean	0.026185	0.024471	0.016155	-0.084057	0.029500
UCL Mean	0.133357	0.131814	0.123648	0.023805	0.136945
Variance	0.987871	0.991014	0.993804	1.000631	0.992911
Stdev	0.993917	0.995497	0.996897	1.000315	0.996449
Skewness	-0.130328	1.888258	-0.156092	-2.677782	0.325238
Excess Kurtosis	5.823385	23.422832	5.651886	143.330987	0.916388
Jarque-Bera statistic	1883.141	31157.97	1775.791	1138420	70.3228
(p-value)	$< 2.2e^{-16}$	$< 2.2e^{-16}$	$< 2.2e^{-16}$	$< 2.2e^{-16}$	$5.551e^{-16}$

Table 9.5 summarizes the descriptive statistics of the residuals to the fitted GARCH(1,1) model. We notice a negative mean value which implies that the residuals of the GARCH(1,1) model were decreasing very slightly. Excess kurtosis is evident which suggests a leptokurtic (heavy tailed) distribution which is sharper with values concentrated around the mean with longer, thicker tails. Negative skewness is visible since skewness = -0.345787. The Jarque-Bera test rejects the hypothesis of Normality at all levels of significance.

9.1.2 Tests for autocorrelation on univariate GARCH(1,1) residuals

After extracting the residuals from the fitted GARCH(1,1) models, we analyze graphically if there exists serial correlation.

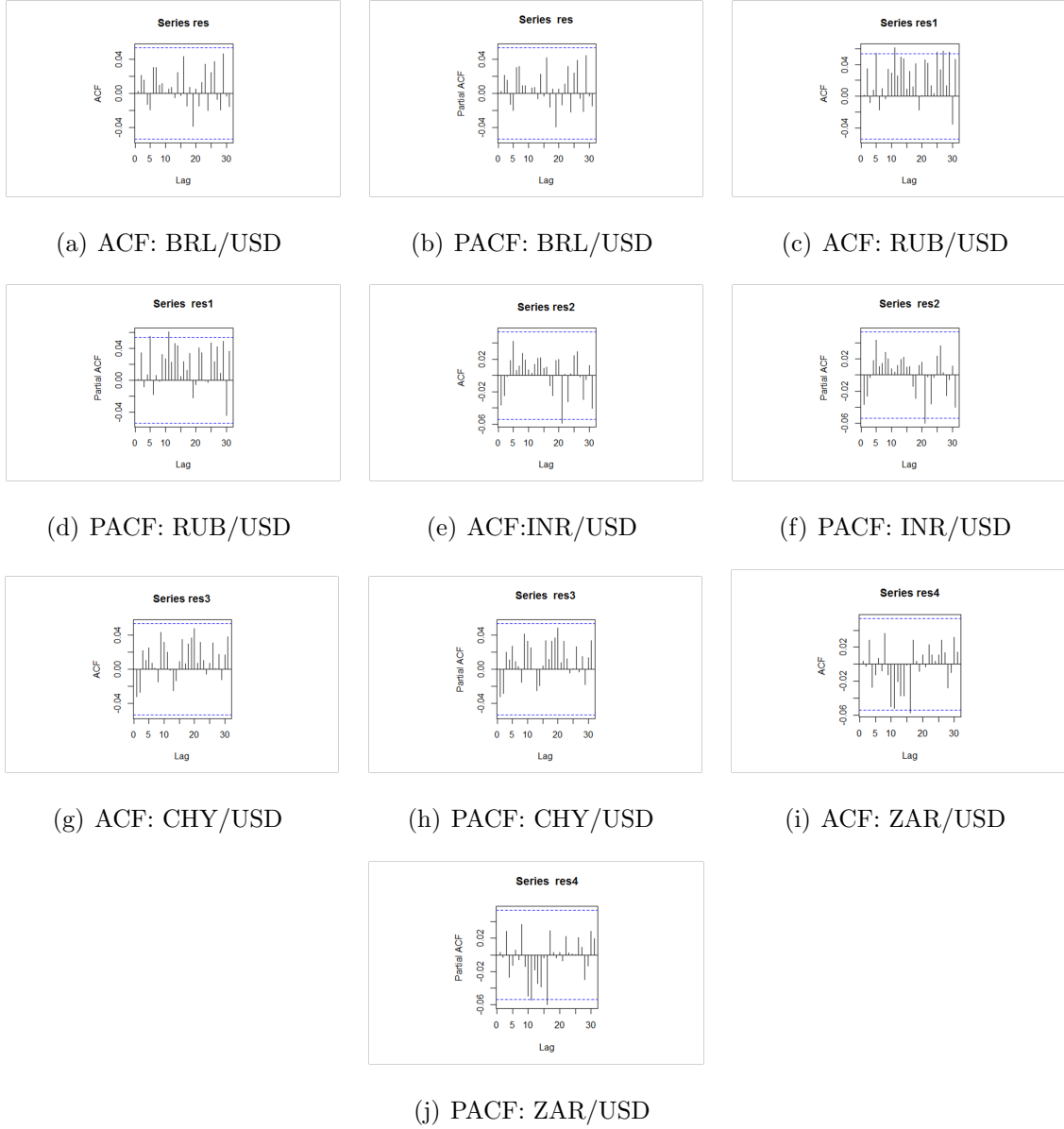


Figure 9.1: ACF and PACF plots of residuals for fitted univariate GARCH(1,1) model.

The ACF and PACF plots in Figure 9.1 show significant lags for the Russian

Ruble, Indian Rupee and South African Rand indicating that the GARCH(1,1) residuals exhibit serial correlation. However, the residuals are independent for the Brazilian Real and Chinese Yuan Renminbi.

Ljung-Box test for autocorrelation:

Table 9.6: Ljung-Box test for autocorrelation for univariate GARCH(1,1) residuals

	Brazil	Russia	India	China	South Africa
Test Statistic	4.4773	9.011	7.243	8.2639	7.8058
<i>p</i> -value	0.9233	0.5311	0.7023	0.6031	0.6478

Table 9.6 shows that at 5% level of significance, we accept the null hypothesis since all *p*-values are greater than 0.05 and conclude that the stable-GARCH(1,1) model for all BRICS countries exchange rates are not serially correlated. The residuals are independent as once the ARCH effects are accounted for, there is evidence to reject the hypothesis of serial correlation.

9.2 Tests for stationarity on univariate GARCH(1,1) residuals

Table 9.7: Stationarity tests for residuals for the fitted univariate GARCH(1,1) model

		Brazil	Russia	India	China	South Africa
ADF	Test Statistic	-10.3584	-9.7371	-10.009	-10.0825	-11.9781
	<i>p</i> -value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
PP	Test Statistic	-1347.08	-1364.054	-1381.667	-1366.924	-1312.903
	<i>p</i> -value	< 0.01	0.01	< 0.01	< 0.01	< 0.01
KPSS	Test Statistic	0.1517	0.7737	0.0498	0.5857	0.082
	<i>p</i> -value	> 0.01	< 0.01	> 0.01	0.02394	> 0.01

We observe for the ADF and P-P tests at 5% and 10% level of significance as previously done, we reject the null hypothesis and conclude that the residuals of GARCH(1,1) model for BRICS to USD exchange rates exhibit stationarity. For the KPSS test, it can be seen that at 5% level of significance, we fail to reject the null hypothesis and conclude that the residuals of GARCH(1,1) model BRICS to USD exchange rates exhibit stationarity.

Subsequently, we fit Nolan's $S_0(\alpha, \beta, \gamma, \delta)$ univariate stable distribution to the residuals.

9.3 Stable parameter estimation

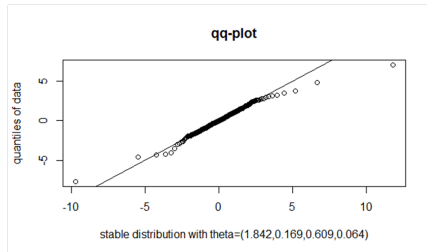
Table 9.8: Stable-GARCH(1,1) parameter estimates for residuals

	Brazil	Russia	India	China	South Africa
α	1.84220452	1.82018772	1.77916146	$1.490059e^{+00}$	1.91177778
β	0.16923594	0.37632959	0.27373087	$1.489665e^{-09}$	0.74065982
γ	0.60891213	0.58295217	0.60095882	$3.569493e^{-01}$	0.67000513
δ	0.06434036	0.01567701	0.02838531	$-2.914503e^{-02}$	0.02663784

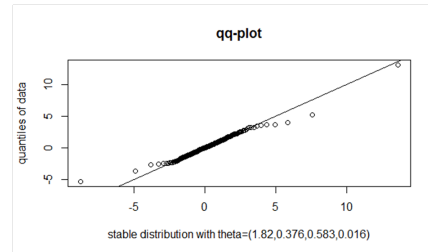
Making a comparison to Table 8.5, we note that the parameters have improved significantly. The index of stability (α) has increased for each country. However, the skewness parameter β still remains relatively close to 0. We also observe a noticeable difference in the scale (γ) and location (δ) parameters. The reason for this difference is that the GARCH filter changes the scale.

9.4 Goodness-of-fit and diagnostics

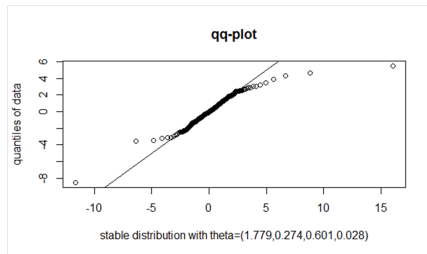
9.4.1 Q-Q plots for stable-GARCH(1,1) residuals



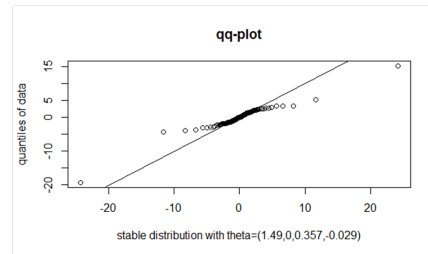
(a) BRL/USD



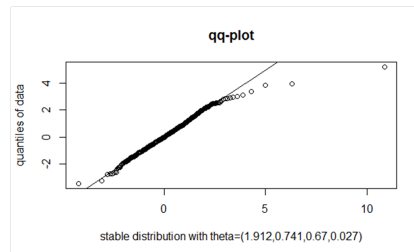
(b) RUB/USD



(c) INR/USD



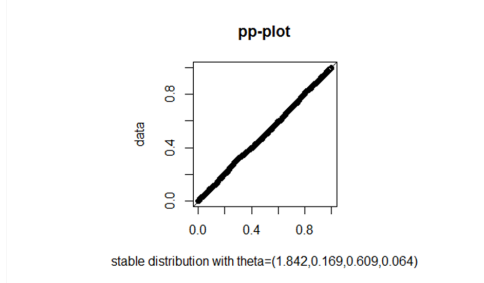
(d) CHY/USD



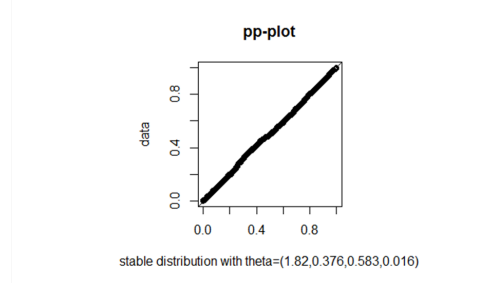
(e) ZAR/USD

Figure 9.2: Q-Q plots for fitted univariate stable-GARCH(1,1) residuals

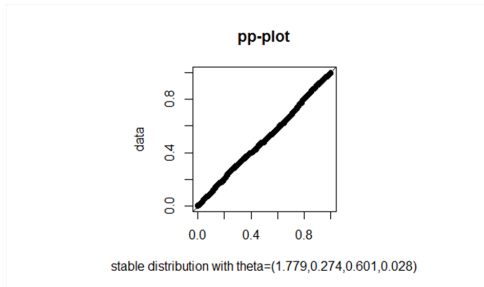
9.4.2 Variance stabilized P-P plots for fitted univariate stable-GARCH(1,1) residuals



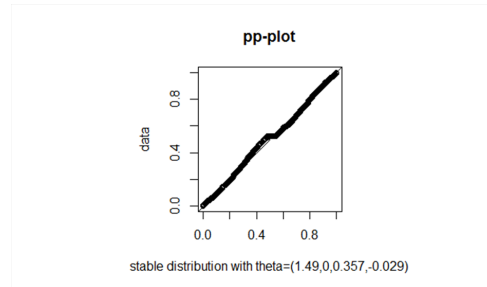
(a) P-P plot of BRL/USD stable-GARCH(1,1) residuals.



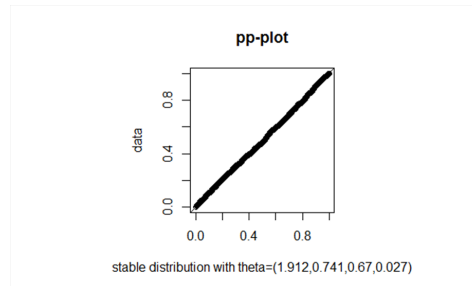
(b) P-P plot of RUB/USD stable-GARCH(1,1) residuals.



(c) P-P plot of INR/USD stable-GARCH(1,1) residuals.



(d) P-P plot of CHY/USD stable-GARCH(1,1) residuals.



(e) P-P plot of ZAR/USD stable-GARCH(1,1) residuals.

Figure 9.3: Variance stabilized P-P plots for fitted univariate stable-GARCH(1,1) residuals

In Figure 9.2, heavy tails are evident in the Q-Q plots and the variance stabilized P-P plots allow for comparisons over the entire range of the residuals.

Although the Q-Q plots imply a poor fit, we focus mainly on the variance stabilized P-P plots for each country in BRICS where a good fit of the residuals to a stable model is evident. The horizontal line segment in subfigure (d) of Figure 9.3 results from the days where the exchange rates remained unchanged on successive days. Overall, the figures graphically show the adequacy of the fitted univariate stable-GARCH(1,1) model *for all* BRICS countries.

		Brazil	Russia	India	China	South Africa
K-S	Test Statistic	0.01923162	0.02398369	0.02531871	0.04385489	0.01912402
	<i>p</i> -value	0.7076644	0.4272273	0.3599577	0.01186553	0.7142049
A-D	Test Statistic	0.3272	0.5493	0.8808	2.1902	0.4432
	<i>p</i> -value	0.9164	0.6972	0.4262	0.07238	0.805

Table 9.9: Goodness-of-fit tests of Stable-GARCH(1,1) residuals

At a 5% level of significance both the K-S and A-D tests indicate that the fitted stable model is adequate in describing the residuals of the stable-GARCH(1,1) model for all countries except China where the K-S test indicates the inadequacy of the stable model at 5% level of significance. At 1% level of significance, the fitted stable model for China is adequate.

9.4.3 stable-GARCH(1,1) density plots

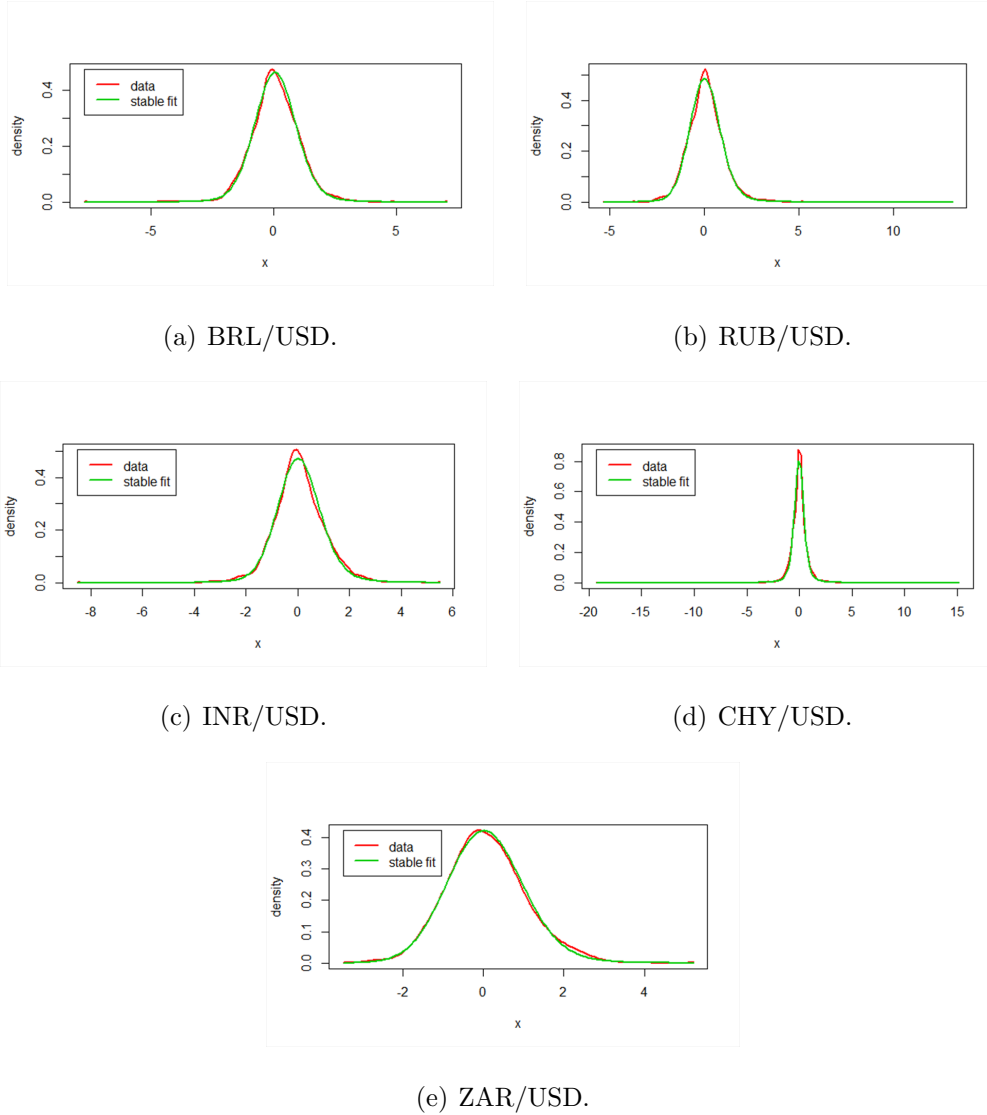


Figure 9.4: Stable density plot of stable-GARCH(1,1) model

The stable density plots indicates that the estimated univariate $S_0(\alpha, \beta, \gamma, \delta)$ model is adequate in describing the residuals extracted through the GARCH(1,1) filter as the residuals slightly deviate from the estimated stable fitted model. As mentioned before, the stable model provides a better fit for the residuals over the entire range where extreme tails are overestimated.

9.5 VaR estimates and backtesting

Table 9.10: VaR estimates of the univariate stable-GARCH(1,1) model

VaR Estimates				
	1%	5%	95%	99%
Brazil	-2.243313	-1.39344	1.602342	2.615574
Russia	-2.094227	-1.335592	1.564488	2.736582
India	-2.337402	-1.399967	1.645003	2.997698
China	-2.850211	-1.128403	1.070113	2.791921
South Africa	-2.133123	-1.477767	1.736747	2.740876

In comparison to the VaR estimates in Table 8.7, we notice that the VaR estimates have improved as shown in Table 9.10 when a GARCH filter is added to the returns.

Table 9.11: Var backtesting for univariate stable-GARCH(1,1) model

	Violation Ratio				<i>p</i> -value of Kupiec Test				<i>p</i> -value of Christoffersen's Test			
	1%	5%	95%	99%	1%	5%	95%	99%	1%	5%	95%	99%
Brazil	0.97	1.04	1.002	1.002	0.73	0.73	0.68	0.52	0.33	0.24	0.35	0.045
Russia	1.2	1.06	0.996	1.001	0.46	0.63	0.55	0.95	0.17	0.71	0.69	0.09
India	1.13	0.89	0.995	1.003	0.63	0.36	0.47	0.12	0.33	0.32	0.05	0.03
China	0.6	1.03	0.994	1.006	0.12	0.82	0.33	0.03	0.12	0.42	0.53	0.01
South Africa	1.06	0.94	0.99	1.002	0.84	0.59	0.28	0.52	0.36	0.48	0.52	0.05

The violation ratios show that the model is a good fit at most VaR levels since most of the ratios comply with the rule of thumb for a precise model: $VR \in [0.8, 1.2]$. Table 9.11 shows that at 1% level of significance, the Kupiec test indicates that the stable model is a good fit for all countries at all VaR levels since all *p*-values are greater than 0.01. The Christoffersen test shows that the fitted stable model is adequate for all countries only at the 5% and 95% VaR levels since the *p*-values are greater than 0.05 at these VaR levels.

Table 9.12: VaR duration test for fitted stable-GARCH(1,1) model

	VaR Duration Test			
	1%		5%	
	Test Statistic	Decision	Test Statistic	Decision
Brazil	0.12	Fail to Reject H_0	0.03	Reject H_0
Russia	0.06	Fail to Reject H_0	0.40	Fail to Reject H_0
India	0.13	Fail to Reject H_0	0.10	Fail to Reject H_0
China	0.74	Fail to Reject H_0	0.43	Fail to Reject H_0
South Africa	0.18	Fail to Reject H_0	0.16	Fail to Reject H_0

The no-hit duration has no-memory hypothesis is accepted. Therefore, we can conclude that the model is correctly specified. We noticed improved results in the stable-GARCH(1,1) case as the GARCH(1,1) filter aids in capturing extreme fluctuations.

9.6 Multivariate stable data analysis

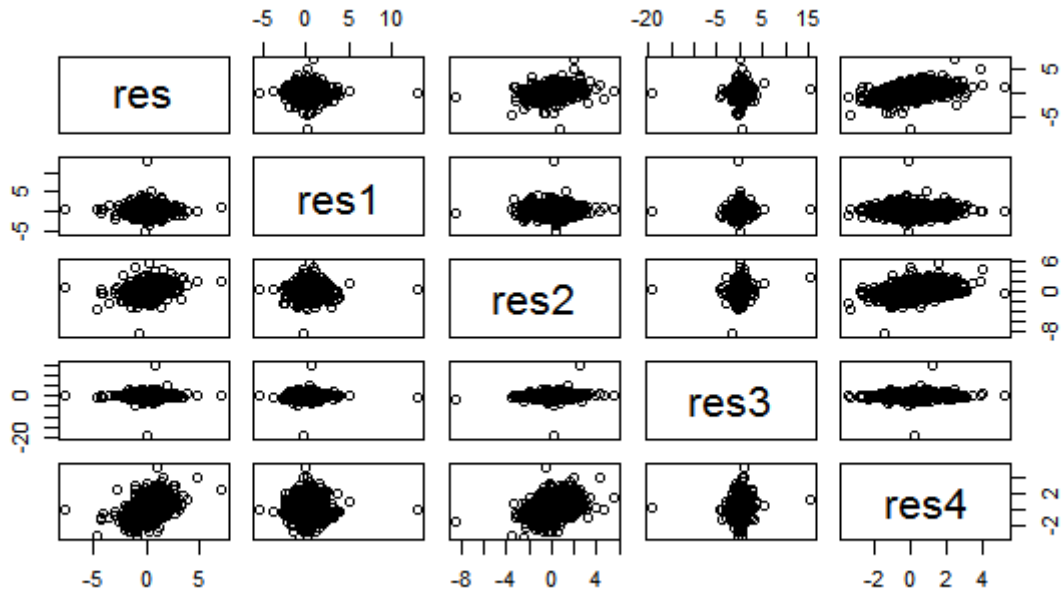


Figure 9.5: Pairwise scatterplots of univariate GARCH(1,1) residuals for BRICS countries to USD exchange rates

The scatterplot matrix shows an approximate elliptical pattern. We proceed to determine the parameters of the stable bivariate elliptical model using the Rachec-Xin-Cheng method of estimation.

Table 9.13: Bivariate elliptical stable parameter estimation for GARCH(1,1) residuals

α	δ	\mathbf{R}
Brazil - Russia		
1.827536	(0.01955281 0.05641651)	$\begin{pmatrix} 0.34606336 & 0.06409438 \\ 0.06409438 & 0.36200815 \end{pmatrix}$
Brazil - India		
1.813726	(0.02920644 0.06401262)	$\begin{pmatrix} 0.37082679 & 0.03083649 \\ 0.03083649 & 0.36405909 \end{pmatrix}$
Brazil - China		
1.690391	(-0.01244109 0.01448774)	$\begin{pmatrix} 0.23451581 & 0.02963903 \\ 0.02963903 & 0.24925545 \end{pmatrix}$
Brazil - South Africa		
1.871113	(0.01444671 0.07260969)	$\begin{pmatrix} 0.4168822 & 0.0539122 \\ 0.0539122 & 0.3965146 \end{pmatrix}$
Russia - India		
1.797345	(0.01547843 0.02846838)	$\begin{pmatrix} 0.33899252 & 0.03585458 \\ 0.03585458 & 0.36093038 \end{pmatrix}$
Russia - China		
1.668664	(-0.01873535 - 0.00938018)	$\begin{pmatrix} 0.21686163 & 0.03454451 \\ 0.03454451 & 0.23337358 \end{pmatrix}$
Russia - South Africa		
1.86538	(-0.001456525 0.028679237)	$\begin{pmatrix} 0.38500077 & 0.04791891 \\ 0.04791891 & 0.39980024 \end{pmatrix}$
India - China		
1.617437	(-0.02259332 - 0.02080865)	$\begin{pmatrix} 2.219397e^{-01} & -4.202606e^{-05} \\ -4.202606e^{-05} & 2.287876e^{-01} \end{pmatrix}$
India - South Africa		
1.85174	(0.01062708 0.04191066)	$\begin{pmatrix} 0.41486610 & 0.02759884 \\ 0.02759884 & 0.39894464 \end{pmatrix}$
China - South Africa		
1.70457	(-0.04080048 - 0.01410422)	$\begin{pmatrix} 0.2627587 & 0.0410898 \\ 0.0410898 & 0.2712789 \end{pmatrix}$

The residuals of the fitted univariate GARCH(1,1) model were fitted with an elliptical stable model with index α , shift δ and shape matrix R . When Q is a multiple of the identity matrix, then the isotropic or radially symmetric case arises. From Table 9.13, there is no substantial empirical evidence for the isotropic heavy tailed stable model for modeling any of the joint pairs of the BRICS countries.

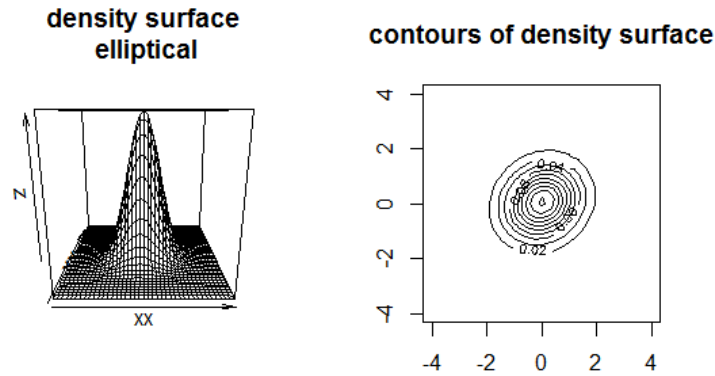


Figure 9.6: Estimated density surface and countour plot for a bivariate elliptical stable fit of Brazillian Real and South African Rand exchange rates of GARCH(1,1) residuals.

The remaining bivariate stable density plots of GARCH(1,1) residuals for BRICS countries may be found in Appendix B.

The fitted spectral measure was used to plot the fitted bivariate density shown in the above figure. The spread of the spectral measure is spiky and masks a pattern that is more evident in the density surface. These are approximately the elliptical countours of the fitted density.

Chapter 10

Conclusion and recommendations

The purpose of this study was to evaluate the fit of stable distributions for BRICS exchange rates. The exchange rate data set spans from the period of January 2011 (when South Africa became a member to the previously known BRIC group) to January 2016. There has been a longstanding debate of whether exchange rate movements should be analyzed in order to comment on a country's financial stability.

This study confirms the results of Nolan (2014) that stable distributions are a flexible class of probability laws that can adequately capture the characteristics of financial data. We have shown that the estimation of stable parameters are feasible and diagnostics prove that large sets of financial data with heavy tails and skewness are well described by stable models as confirmed by the Kolmogorov-Smirnov and Anderson-Darling tests.

VaR estimates and VaR in-sample backtesting using violation ratios, the Kupiec likelihood ratio test, Christoffersen conditional coverage test and the VaR duration test emphasize the robustness of the fitted stable models. The use

of stable distributions for data in finance is largely justified in this study by capturing large fluctuations which is frequently seen in the financial industry where the need for better models is imperative for acknowledging the many empirical properties of financial data. Policy makers, regulators, risk adverse investors and insurers could leverage the most by using stable distributions as they are parties that remain largely concerned about extreme losses.

This research recommends further study in:

- The use of Value-at-Risk (VaR) estimates and backtesting to evaluate the performance in the multivariate stable case.
- Comparisons to other distributions such as the generalized hyperbolic distribution or generalized lambda distribution.
- Possible alternative criteria for evaluation and model selection based on the tails of the data.

References

- Abdul-Hamid, H. (1996). *Approximation of Multivariate stable densities*. Ph.D. thesis, American University.
- Araujo, A., & Giné, E. (1980). *The Central Limit Theorem for Real and Banach valued random variables*, vol. 431. Wiley New York.
- Arezki, R., Dumitrescu, E., & Quintyn, M. (2012). Commodity prices and exchange rate volatility: Lessons from south africa’s capital account liberalization.
- BiiiCPA (2016). The basel iii accord.
URL [http://www.basel-iii-accord.com/Accessed\(11November2016\)](http://www.basel-iii-accord.com/Accessed(11November2016))
- Blanco, C., & Oks, M. (2004). Backtesting var models: Quantitative and qualitative tests. *The Risk Desk*, 4(1).
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3), 307–327.
- Buckle, D. (1995). Bayesian inference for stable distributions. *Journal of the American Statistical Association*, 90(430), 605–613.
- Byczkowski, T., Nolan, J. P., & Rajput, B. (1993). Approximation of multidimensional stable densities. *Journal of Multivariate Analysis*, 46(1), 13–31.

- Campa, J. M., Chang, P. K., & Refalo, J. F. (2002). An options-based analysis of emerging market exchange rate expectations: Brazil's real plan, 1994–1999. *Journal of Development Economics*, 69(1), 227–253.
- Campbell, S. D. (2006). A review of backtesting and backtesting procedures. *The Journal of Risk*, 9(2), 1.
- Caporale, G. M., Spagnolo, F., & Spagnolo, N. (2016). Macro news and exchange rates in the brics.
- Cheng, B., & Rachev, S. T. (1995). Multivariate stable futures prices. *Mathematical Finance*, 5(2), 133–153.
- Chinhamu, K., Huang, C.-K., & Chikobvu, D. (2015). Evaluating risk in gold prices with generalized hyperbolic and stable distributions. In *Proceedings of the 57th Annual Conference of SASA*, vol. 17, (p. 24).
- Christoffersen, P., & Pelletier, D. (2004). Backtesting value-at-risk: A duration-based approach. *Journal of Financial Econometrics*, 2(1), 84–108.
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, (pp. 841–862).
- Daniélsson, J. (2011). Financial risk forecasting: the theory and practice of forecasting market risk with implementation in r and matlab.
- Dasgupta, R. (2014). Integration and dynamic linkages of the indian stock market with bric-an empirical study. *Asian Economic and Financial Review*, 4(6), 715.
- de Jager, S. (2012). Working paper-wp/12/02-modelling south africa's equilibrium real effective exchange rate-a vecm approach. Tech. rep., South African Reserve Bank.
- DuMouchel, W. (1971). *Stable distributions in statistical inference, unpublished Ph. D.* Ph.D. thesis, thesis, Department of Statistics, Yale University.

- DuMouchel, W. H. (1973). On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. *The Annals of Statistics*, (pp. 948–957).
- Embrechts, P., Resnick, S. I., & Samorodnitsky, G. (1999). Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3(2), 30–41.
- eNCA (2016). Rand plummets overnight to record low before recovering.
 URL [https://www.enca.com/south-africa/
 rand-hits-record-low-one-night](https://www.enca.com/south-africa/rand-hits-record-low-one-night) Accessed (05 June 2016)
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, (pp. 987–1007).
- Evans, D. L., Drew, J. H., & Leemis, L. M. (2008). The distribution of the kolmogorov–smirnov, cramer–von mises, and anderson–darling test statistics for exponential populations with estimated parameters. *Communications in Statistics—Simulation and Computation*®, 37(7), 1396–1421.
- Fama, E. F., & Roll, R. (1968). Some properties of symmetric stable distributions. *Journal of the American Statistical Association*, 63(323), 817–836.
- Feldheim, E. (1937). Etude de la stabilité des lois de probabilité. *Thèses françaises de l'entre-deux-guerres*, 187, 1–64.
- Fisher, R. A. (1929). Tests of significance in harmonic analysis. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 125(796), 54–59.
- Greenspan, A. (1996). Remarks at the financial markets conference of the federal reserve bank of atlanta.
- Gujarati, D. (2004). Basic econometrics.

- Ijumba, C. (2013). Multivariate analysis of the brics financial markets.
- Jorion, P. (2007). *Value at Risk: The New Benchmark for Managing Financial Risk*. New York: McGraw-Hill.
- Katsenga, G. Z. (2013). *Value at Risk (VaR) Backtesting: 'evidence from a South African Market Portfolio'*.
- Kemda, L. E., Huang, C.-K., & Chinhamu, K. (2015). Value-at-risk for the usd/zar exchange rate: The variance-gamma model. *South African Journal of Economic and Management Sciences*, 18(4), 551–566.
- Kogon, S. M., & Williams, D. B. (1998). Characteristic function based estimation of stable distribution parameters. *A practical guide to heavy tails: statistical techniques and applications*, (pp. 311–338).
- Koutrouvelis, I. A. (1980). Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, 75(372), 918–928.
- Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The J. of Derivatives*, 3(2).
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3), 159–178.
- Ločmelis, K., Mititel, D., Lubláy, Á., & Ajevskis, V. (2015). Impact of russiaâs 2014–2015 crisis on the dynamic linkages between the stock markets of russia, the eu and us.
- Ma, M., Du, J., & Lai, K. K. (2013). Modeling volatility of exchange rate of chinese yuan against us dollar based on garch models. In *Business Intelligence and Financial Engineering (BIFE), 2013 Sixth International Conference on*, (pp. 295–299). IEEE.

- Mandalos, A. T. (2014). *A time series analysis of the ZAR/USD exchange rate..* Ph.D. thesis, University of Macedonia.
- McCulloch, J. (1994). Estimation of bivariate stable spectral densities. *Unpublished manuscript, Department of Economics, Ohio State University.*
- McCulloch, J. H. (1986). Simple consistent estimators of stable distribution parameters. *Communications in Statistics-Simulation and Computation*, 15(4), 1109–1136.
- McCulloch, J. H. (1997). Measuring tail thickness to estimate the stable index α : A critique. *Journal of Business & Economic Statistics*, 15(1), 74–81.
- McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management: Concepts, techniques and tools*. Princeton university press.
- Murari, K., & Sharma, R. (2013). Ols modeling for indian rupee fluctuations against us dollar. *Global Advanced Research Journal of Management and Business Studies (ISSN: 2315-5086) Vol, 2(12)*, 559–566.
- Nelson, R. M. (2013). Current debates over exchange rates: Overview and issues for congress.
- Nguyen, T. T., & Sampson, A. R. (1991). A note on characterizations of multivariate stable distributions. *Annals of the Institute of Statistical Mathematics*, 43(4), 793–801.
- Nikias, C. L., & Shao, M. (1995). *Signal processing with alpha-stable distributions and applications*. Wiley-Interscience.
- NIST/SEMATECH (2013). 1.3.3.26.11. scatterplot matrix.
URL <http://itl.nist.gov/div898/handbook/eda/section3/eda33qb.htm>
- Nolan, J. P. (2001). Maximum likelihood estimation and diagnostics for stable distributions. In *Lévy processes*, (pp. 379–400). Springer.

- Nolan, J. P. (2003). Modeling financial data with stable distributions. *Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance: Book, 1*, 105–130.
- Nolan, J. P. (2005). Multivariate stable densities and distribution functions: general and elliptical case. In *Deutsche Bundesbank's 2005 annual fall conference*.
- Nolan, J. P. (2006). Multivariate elliptically contoured stable distributions: Theory and estimation. *Citeseer*.
- Nolan, J. P. (2008). An overview of multivariate stable distributions. *Online: <http://academic2.american.edu/~jpnolan/stable/overview.pdf>*.
- Nolan, J. P. (2014). Financial modeling with heavy-tailed stable distributions. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(1), 45–55.
- Nolan, J. P. (2015). *Stable Distributions - Models for Heavy Tailed Data*. Boston: Birkhauser. In progress, Chapter 1 online at academic2.american.edu/~jpnolan.
- Nolan, J. P., Panorska, A. K., & McCulloch, J. H. (2001). Estimation of stable spectral measures. *Mathematical and Computer Modelling*, 34(9), 1113–1122.
- Ojeda, D. (2001). *Comparison of stable estimators*. Ph.D. thesis.
- Oral, E., Erdemir, C., et al. (2012). A bayesian estimation of stable distributions. *Journal of Statistical and Econometric Methods*, 1(3), 39–52.
- Picardo, C. E. (2014). Exchange rate.
URL <http://www.investopedia.com/terms/e/exchangerate.asp> Accessed(9July2016)
- Poon, S.-H. (2008). Arch.
URL <https://phps.portals.mbs.ac.uk/Portals/49/docs/spoon/ARCH.pdf.pdf>

- Pradhan, R. P., Dasgupta, P., & Samadhan, B. (2013). Finance development and economic growth in brics: a panel data analysis. *Journal of Quantitative Economics*, 11(1-2), 308–322.
- Press, S. J. (1972). Multivariate stable distributions. *Journal of Multivariate Analysis*, 2(4), 444–462.
- Rachev, S. T., & Xin, H. (1993). Test for association of random variables in the domain of attraction of multivariate stable law. *Probability and mathematical statistics*, 14(1), 125–141.
- Samoradnitsky, G., & Taqqu, M. S. (1994). *Stable non-Gaussian random processes: stochastic models with infinite variance*, vol. 1. CRC press.
- SARB (2015). South africaâs implementation of basel ii and basel iii.
URL [https://www.resbank.co.za/RegulationAndSupervision/BankSupervisionAccessed\(23May2016\)](https://www.resbank.co.za/RegulationAndSupervision/BankSupervisionAccessed(23May2016))
- Schuster, A. (1898). On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Terrestrial Magnetism*, 3(1), 13–41.
- Shumway, R. H., & Stoffer, D. S. (2010). *Time series analysis and its applications: with R examples*. Springer Science & Business Media.
- StataCorp (2015). User’s guide-stata.
URL <http://www.stata.com/manuals14/u.pdf>Accessed(16July2016)
- Stoyanov, S. V., Rachev, S. T., Racheva-Yotova, B., & Fabozzi, F. J. (2011). Fat-tailed models for risk estimation. *Journal of Portfolio Management*, 37(2), 107.
- Sule, A. (2011). Brics can build common currency.
URL http://europe.chinadaily.com.cn/epaper/2011-04/08/content_12291921.htm

- Syczewska, E. M., et al. (2010). Empirical power of the kwiatkowski-phillips-schmidt-shin test. Tech. rep.
- Tsay, R. S. (2005). *Analysis of financial time series*, vol. 543. John Wiley & Sons.
- Tsay, R. S. (2013). *Multivariate Time Series Analysis: with R and financial applications*. John Wiley & Sons.
- Uchaikin, V. V., & Zolotarev, V. M. (1999). *Chance and stability: stable distributions and their applications*. Walter de Gruyter.
- Walters, S., & De Beer, B. (1999). An indicator of south africaâs external competitiveness. *South African Reserve Bank Quarterly Bulletin*, 213, 54–67.
- Wei, W. (2006). *Time Series Analysis: univariate and multivariate methods*. Boston: Pearson Addison Wesley.
- Yang, G. (2012a). *The energy goodness-of-fit test for univariate stable distributions*. Ph.D. thesis, Bowling Green State University.
- Yang, Y. (2012b). Parameter estimation of garch model. *Optimization*, 1, 1.
- Zolotarev, V. M. (1986). *One-dimensional stable distributions*, vol. 65. American Mathematical Soc.

Appendix

Appendix A

A.1. Generalized Central limit theorem

The Central limit Theorem (CLT) states that the sum of independent and identically distributed random variables converges to a normal distribution when the sum is centered and scaled appropriately and the number of terms within the summation tend to go to infinity. The Generalized Central limit theorem assumes that random variables with infinite variance also converge, but to a stable distribution not necessarily a normal distribution

Generalized Central Limit Theorem

A non-degenerate random variable Y is α - *stable* for some $0 < \alpha \leq 2$ if and only if there is an independent and identically distributed sequence of random variables $X_1, X_2, X_3, ..$ and constants $a_n > 0, b_n \in \mathbb{R}$ with

$$a_n(X_1 + \dots + X_n) - b_n \xrightarrow{d} Y.$$

A.2. Conversions Between parameterizations

$S_1 \rightarrow S_0$

$$\beta_0 = \beta_1, \gamma_0 = \gamma_1, \delta_0 = \begin{cases} \delta_1 + \beta\gamma \tan \frac{\pi\alpha}{2}; & \alpha \neq 1, \\ \delta_1 - \beta \frac{2}{\pi} \gamma \ln \gamma & \alpha = 1. \end{cases}$$

$$(B) \rightarrow (A)$$

$$\beta_A = \beta_B, \delta_A = \frac{2}{\pi}\delta_B, \gamma_A = \frac{\pi}{2}\gamma_B \quad \text{if } \alpha = 1;$$

$$\begin{cases} \beta_A = \cot \frac{\pi}{2} \alpha \tan(\frac{\pi}{2} \beta_B K(\alpha)), \\ \delta_A = \frac{\delta_B}{\cos(\frac{\pi}{2} \beta_B K(\alpha))}, \\ \gamma_A = \gamma_B \cos(\frac{\pi}{2} \beta_B K(\alpha)) \end{cases} \quad \text{if } \alpha \neq 1,$$

$$(A) \rightarrow (B)$$

$$\beta_B = \beta_A, \delta_B = \frac{2}{\pi}\delta_A, \gamma_B = \frac{\pi}{2}\gamma_A \quad \text{if } \alpha = 1;$$

$$\begin{cases} \beta_B = \frac{2}{\pi K(\alpha)} \left(\arctan \frac{\beta_A}{\cos \frac{\pi}{2} \alpha} \right), \\ \delta_B = \delta_A \left(\frac{\cos^2(\frac{\pi}{2} \alpha)}{\beta_A^2 + \cos^2 \frac{\pi}{2} \alpha} \right)^{\frac{1}{2}}, \\ \gamma_B = \gamma_A \left(\frac{\cos^2(\frac{\pi}{2} \alpha)}{\beta_A^2 + \cos^2 \frac{\pi}{2} \alpha} \right)^{-\frac{1}{2}}. \end{cases} \quad \text{if } \alpha \neq 1$$

$$(A) \rightarrow (M)$$

$$\begin{aligned} \beta_M &= \beta_A, \delta_M = \delta_A, \gamma_M = \gamma_A && \text{if } \alpha = 1; \\ \beta_M &= \beta_A, \delta_M = \delta_A + \beta_A \tan \frac{\pi \alpha}{2}, \gamma_M = \gamma_A && \text{if } \alpha \neq 1, \end{aligned}$$

$$(A) \rightarrow S_1$$

$$\beta_1 = \beta_A, \delta_1 = \gamma_A \delta_A, \gamma_1^\alpha = \gamma_A$$

$$S_1 \rightarrow (A)$$

$$\beta_A = \beta_A, \gamma_A = \gamma_1^\alpha, \delta_A = \frac{\delta_1}{\gamma_1^\alpha}$$

(Yang, 2012a)

$$(B) \rightarrow (C)$$

$$\begin{aligned} \theta &= \frac{2}{\pi} \arctan(\frac{2\gamma_B}{\pi}), \delta_C = \delta_B(\frac{\pi^2}{4} + \gamma_B^2)^{1/2} & \text{if } \alpha = 1; \\ \theta &= \beta_B K(\alpha)/\alpha, \delta_C = \delta_B & \text{if } \alpha \neq 1. \end{aligned}$$

$$(C) \rightarrow (E)$$

$$\nu = \alpha^2, \theta_E = \theta_C, \tau = \frac{1}{\alpha} \log \delta_C + \mathbb{C} \left(\frac{1}{\alpha - 1} \right) \text{ (Zolotarev, 1986)}.$$

A.3. Isotropic stable distributions The isotropic or radially symmetric case arises when Σ (some positive definite matrix) is a multiple of the identity matrix leading to the simplification of the following characteristic equation as:

$$E \left[e^{i\mathbf{u}^T \mathbf{Y}} \right] = e^{-\gamma_0^\alpha |\mathbf{u}|^\alpha + i\mathbf{u}^T \delta} \quad (10.1)$$

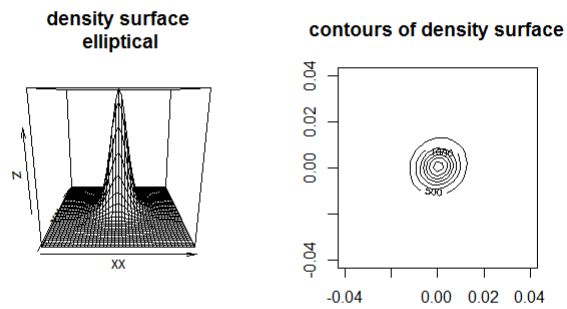
and projection parameter functions;

$$\gamma(\mathbf{u}) = \gamma_0 |\mathbf{u}|, \quad \beta(\mathbf{u}) = 0, \quad \delta(\mathbf{u}) = \langle \mathbf{u}, \delta \rangle$$

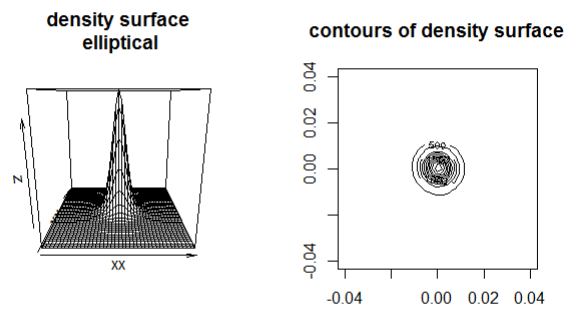
where γ_0 is a scale parameter and $\delta \in \mathbb{R}^d$ is a location parameter. The spectral measure in this case is a uniform distributon on a unit sphere $\mathbb{S} = \{\mathbf{y}^T \mathbf{y} = 1\} \subset \mathbb{R}^d$, where $A \sim S(\frac{\alpha}{2}, 1, 2\gamma_0^2(\cos(\frac{\pi\alpha}{4})^{\frac{2}{\alpha}}), 0)$ and $G \sim N(0, \mathbf{I})$, then $\mathbf{Y} = A^{\frac{1}{2}} \mathbf{G} + \delta$.

Appendix B

Estimated density surface and countour plots for a bivariate elliptical stable fit of pairwise BRICS/USD countries

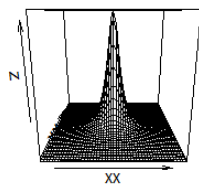


(a) BRL-RUB/USD

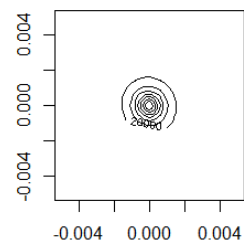


(b) BRL-INR/USD

**density surface
elliptical**

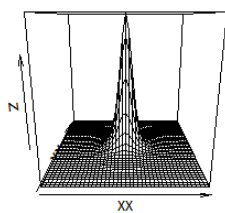


contours of density surface

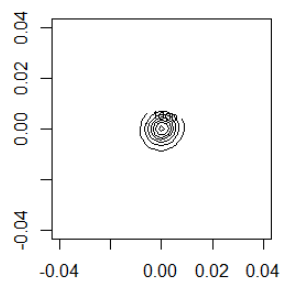


(c) BRL-CHY/USD

**density surface
elliptical**

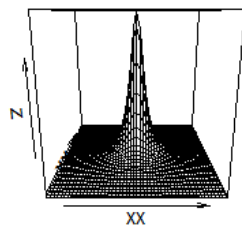


contours of density surface

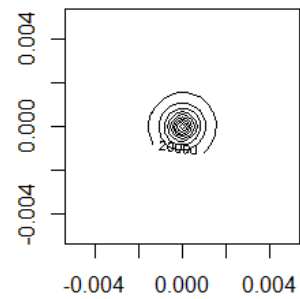


(d) RUB-INR/USD

**density surface
elliptical**

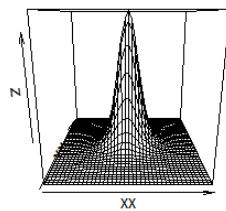


contours of density surface

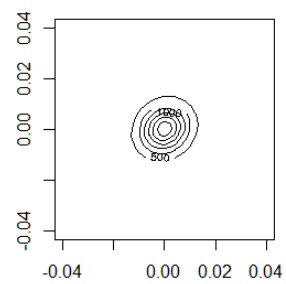


(e) RUB-INR/USD

**density surface
elliptical**

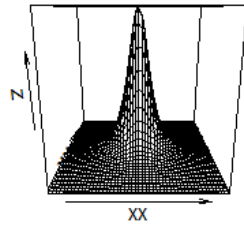


contours of density surface

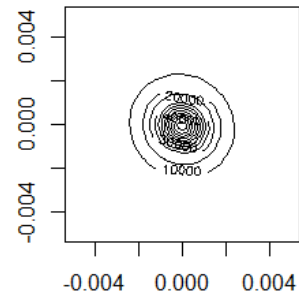


(f) RUB-ZAR/USD

**density surface
elliptical**

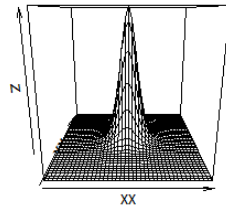


contours of density surface

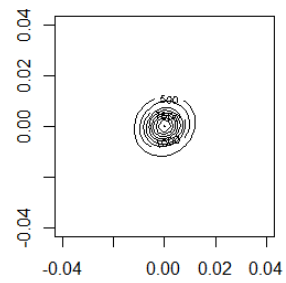


(g) INR-CHY/USD

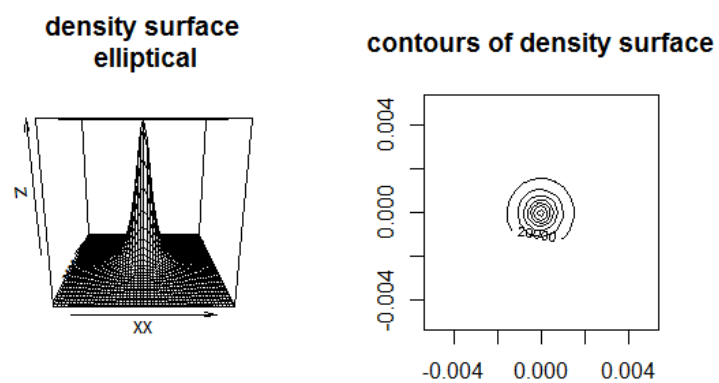
**density surface
elliptical**



contours of density surface



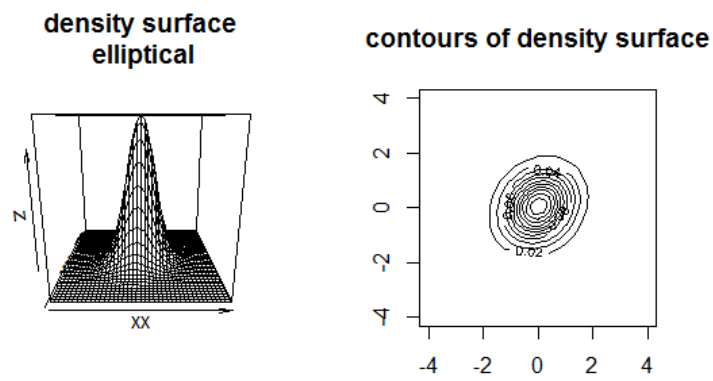
(h) INR-ZAR/USD



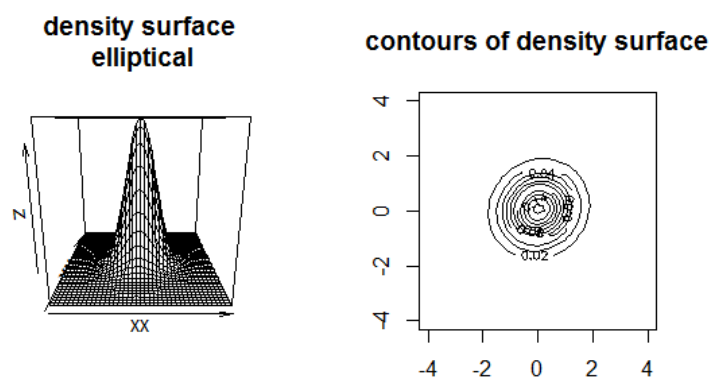
(i) CHY-ZAR/USD

Figure 10.1: Estimated density surfaces and contour plots for a bivariate elliptical stable fit of BRICS countries exchange rates.

Combining stable and GARCH(1,1) model: Estimated density surface and countour plots

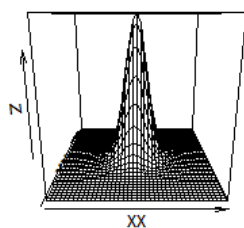


(a) BRL-RUB/USD

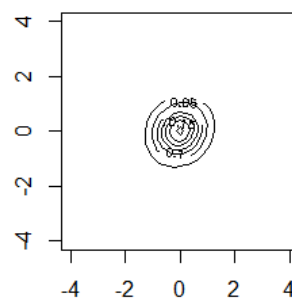


(b) BRL-INR/USD

**density surface
elliptical**

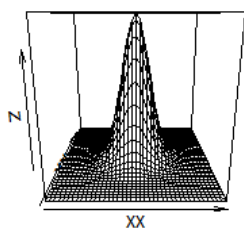


contours of density surface

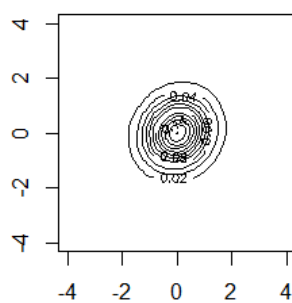


(c) BRL-CHY/USD

**density surface
elliptical**

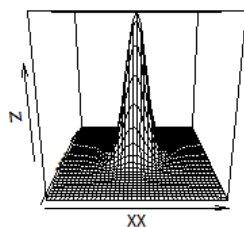


contours of density surface

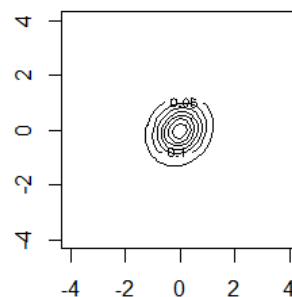


(d) RUB-INR/USD

**density surface
elliptical**

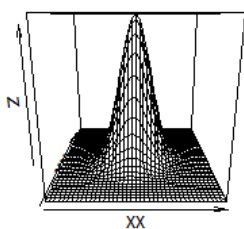


contours of density surface

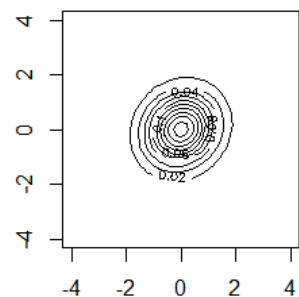


(e) RUB-CHY/USD

**density surface
elliptical**

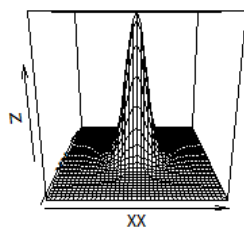


contours of density surface

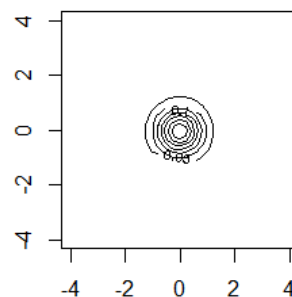


(f) RUB-ZAR/USD

**density surface
elliptical**

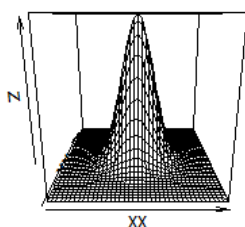


contours of density surface

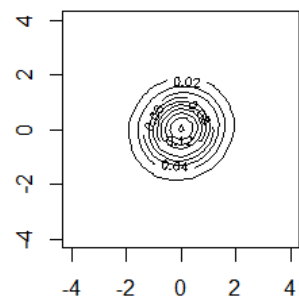


(g) INR-CHY/USD

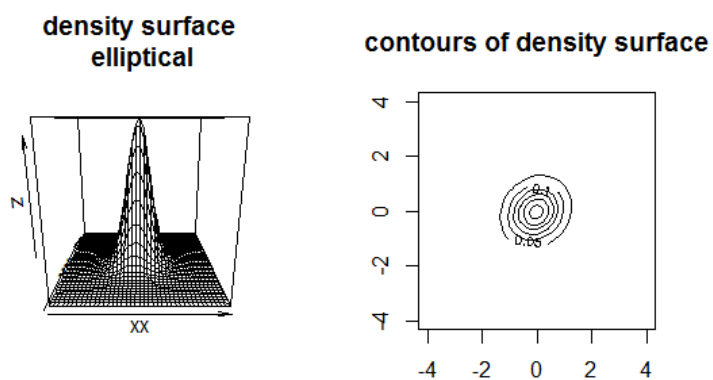
**density surface
elliptical**



contours of density surface



(h) INR-ZAR/USD



(i) CHY-ZAR/USD

Figure 10.2: Bivariate elliptical stable density plots of GARCH(1,1) residuals for BRICS countries

R code

```
> View(BRICS.USD)
> BRL<-(BRICS.USD)[,c("V2")]
> BRL1= diff(log(BRL))
> RUB<-(BRICS.USD)[,c("V2")]
> RUB1=diff(log(RUB))
> INR<-(BRICS.USD)[,c("V2")]
> INR1=diff(log(INR))
> CHY<-(BRICS.USD)[,c("V2")]
> CHY1=diff(log(CHY))
> ZAR<-(BRICS.USD)[,c("V2")]
> ZAR1=diff(log(ZAR))
```

The R code provided below yields empirical results for Brazil, unless where labelled/defined otherwise. Similar codes are used for the remaining BRICS countries with their respective parameter estimates by adjusting the R code accordingly.

Time series plots

```
> code=ts(BRL,frequency=260,start=c(2011,140))
> plot(code,xlab='Year',ylab='Real/USD',main='Time series plot of Real/USD')
> code=ts(BRL1,frequency=260,start=c(2011,140))
> plot(code,xlab='Year',ylab='Real/USD returns',main='Time series plot of Real/USD returns')
```

Descriptive statistics

```
> basicStats(BRL1)
> jarqueberaTest(BRL1)
```

ACF and PACF plots

```
> acf(BRL1)
> pacf(BRL1)
```

Ljung-Box test for autocorrelation

```
> Box.test(BRL1, lag = 10, type = c("Ljung-Box"))
```

Tests for stationarity

ADF test

```
> adf.test(BRL1)
```

P-P test

```
> pp.test(BRL1)
```

KPSS test

```
> kpss.test(BRL1)
```

Fitted stable distribution

```
> stable.fit(BRL1, method = 1, param = 0)
```

Goodness of fit tests and diagnostics

```
> theta=c(1.736639e+00,-2.102931e-08,5.323732e-03,6.868068e-04)
```

```
> stable.ks.gof(BRL1,theta,method=0,param=0)
```

```
> ad.test(BRL1,pstable,alpha= 1.736639e+00,beta= -2.102931e-08,gamma= 5.323732e-03,delta=
6.868068e-04)
```

```
> qqstable(BRL1, theta, param = 0, ptwise.ci = FALSE)
```

```
> ppstable(BRL1, theta, var.stabilized = FALSE, param = 0)
```

```
> stable.density.plot(BRL1, theta, param=0, xrange = range(BRL1))
```

VaR and in-sample backtesting

```
> qstable(0.01, 1.736639e+00, -2.102931e-08 ,5.323732e-03 ,6.868068e-04 ,0)
```

```
> qstable(0.05, 1.736639e+00, -2.102931e-08 ,5.323732e-03 ,6.868068e-04 ,0)
```

```
> qstable(0.95, 1.736639e+00, -2.102931e-08 ,5.323732e-03 ,6.868068e-04 ,0)
```

```
> qstable(0.99, 1.736639e+00, -2.102931e-08 ,5.323732e-03 ,6.868068e-04 ,0)
```

```
> VaRTest(0.01,BRL1,rep(-0.02487615,1324))
```

```
VaRTest(0.05,BRL1,rep(-0.01307162,1324))
```

```
> VaRTest(0.95,BRL1,rep(0.01444523,1324))
```

```
> VaRTest(0.99,BRL1,rep(0.02624979,1324))
```

```
> VaRDurTest(0.01,BRL1,rep(-0.02487615,1324),conf.level = 0.95)
```

```
> VaRDurTest(0.05,BRL1,rep(-0.01307162,1324),conf.level = 0.95)
```

```
> VaRDurTest(0.95,BRL1,rep(0.01444523,1324),conf.level = 0.95)
> VaRDurTest(0.99,BRL1,rep(0.02624979,1324),conf.level = 0.95)
```

Scatterplot matrix

```
> View(BRICS.USD)
> B<-(BRICS.USD)[,c("V2")]
> R<-(BRICS.USD)[,c("V3")]
> I<-(BRICS.USD)[,c("V4")]
> C<-(BRICS.USD)[,c("V5")]
> S<-(BRICS.USD)[,c("V6")]
> Brazil=diff(log(B))
> Russia=diff(log(R))
> India=diff(log(I))
> China=diff(log(C))
> SouthAfrica=diff(log(S))
> pairs( Brazil+Russia+India+China+SouthAfrica)
```

Multivariate elliptical stable parameter estimation

R code for 10 pairwise combinations of BRICS countries:

```
> BR<-matrix(c(BRL1,RUB1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(BR, method1d=1)
> BI<-matrix(c(BRL1,INR1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(BI, method1d=1)
> BC<-matrix(c(BRL1,CHY1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(BC, method1d=1)
> BZ<-matrix(c(BRL1,ZAR1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(BZ, method1d=1)
> RI<-matrix(c(RUB1,INR1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(RI, method1d=1)
> RC<-matrix(c(RUB1,CHY1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(RC, method1d=1)
> RZ<-matrix(c(RUB1,ZAR1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(RZ, method1d=1)
> IC<-matrix(c(INR1,CHY1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(IC, method1d=1)
> IZ<-matrix(c(INR1,ZAR1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(IZ, method1d=1)
> CZ<-matrix(c(CHY1,ZAR1),nrow = 2,ncol = 1324)
> mvstable.fit.elliptical(CZ, method1d=1)
```

Estimated density and contour plots

```
> stable.test ( xx <- seq(-0.04, 0.04, 0.002)
d.ell <- mvstable.elliptical(alpha = 1.523953, R = matrix(c(2.422102e-05, 2.301749e-06,
2.301749e-06, 2.482960e-05 ), 2, 2), delta = c(0.0002700612,0.0005117953))
pdf.surface.plots(d.ell, xx, "elliptical"))
```


Combining stable and GARCH(1,1) model

```
> BRL<-(BRICS.USD)[c("V2")]
```

```
> BRL1=diff(log(BRL))
```

ARCH LM test

```
> ArchTest(BRL1,lag=12,demean = FALSE)
```

Fitting GARCH(1,1) model

```
> m1=garchFit(formula= garch(1,1),data=BRL1,trace=F)
```

```
> summary(m1)
```

```
> m2=garchFit(formula= garch(1,1),data=BRL1,include.mean=FALSE,trace=F)
```

```
> summary(m2)
```

```
> res=residuals(m2,standardize=TRUE)
```

Descriptive statistics

```
> basicStats(res)
```

ACF and PACF plots

```
> acf(res)
```

```
> pacf(res)
```

Ljung-Box test for autocorrelation

```
> Box.test(res, lag = 10, type = c("Ljung-Box"))
```

Tests for stationarity

```
> adf.test(res)
```

```
> pp.test(res)
```

```
> kpss.test(res)
```

Stable parameter estimation

```
> stable.fit(res,method = 1,param = 0)
```

```
> theta=c(1.84220452,0.16923594,0.60891213,0.06434036)
```

Goodness-of-fit tests and diagnostics

```
> stable.ks.gof(res, theta, method=1, param=0)
```

```
> ad.test(res,pstable,alpha=1.84220452,beta=0.16923594,gamma=0.60891213,delta=0.064340)
```

```
> stable.density.plot(res, theta, param=0, xrange = range(res))
```

```
> qqstable(res, theta, param = 0, ptwise.ci = FALSE)
```

```
> ppstable(res, theta, var.stabilized = FALSE, param = 0)
```

VaR and backtesting

```
> qstable(0.01,1.84220452,0.16923594,0.60891213,0.06434036)
```

```
> qstable(0.05,1.84220452,0.16923594,0.60891213,0.06434036)
```

```
> qstable(0.95,1.84220452,0.16923594,0.60891213,0.06434036)
```

```
> qstable(0.99,1.84220452,0.16923594,0.60891213,0.06434036)
```

```
> VaRTest(0.01,res,rep(-2.243313,1324))
```

```
> VaRTest(0.05,res,rep(-1.39344,1324))
```

```
> VaRTest(0.95,res,rep(1.602342,1324))
```

```
> VaRTest(0.99,res,rep(2.615574,1324))
```

```
> VaRDurTest(0.01,res,rep(-2.243313,1324),conf.level = 0.95)
```

```
> VaRDurTest(0.05,res,rep(-1.39344,1324),conf.level = 0.95)
```

```
> VaRDurTest(0.95,res,rep(1.602342,1324),conf.level = 0.95)
```

```
> VaRDurTest(0.99,res,rep(2.615574,1324),conf.level = 0.95)
```

Scatterplot matrix

```
> BRL<-(BRICS.USD)[,c("V2")]
```

```
> BRL1=diff(log(BRL))
```

```
> m2=garchFit(formula= garch(1,1),data=BRL1,include.mean=FALSE,trace=F)
```

```
> summary(m2)
```

```
> m2=garchFit(formula= garch(1,1),data=BRL1,include.mean=FALSE,trace=F)
```

```
> res=residuals(m2,standardize=TRUE)
```

```
> RUB<-(BRICS.USD)[,c("V3")]
```

```
> RUB1=diff(log(RUB))
```

```
> m4=garchFit(formula= garch(1,1),data=RUB1,include.mean=FALSE,trace=F)
```

```
> res1=residuals(m4,standardize=TRUE)
```

```
> INR<-(BRICS.USD)[,c("V4")]
```

```
> INR1=diff(log(INR))
```

```
> m6=garchFit(formula= garch(1,1),data=INR1,include.mean=FALSE,trace=F)
```

```
> res2=residuals(m6,standardize=TRUE)
```

```
> CHY<-(BRICS.USD)[,c("V5")]
```

```

> CHY1=diff(log(CHY))
> m8=garchFit(formula= garch(1,1),data=CHY1,include.mean=FALSE,trace=F)
> res3=residuals(m8,standardize=TRUE)
> ZAR<-(BRICS.USD)[,c("V6")]
> ZAR1=diff(log(ZAR))
> m10=garchFit(formula= garch(1,1),data=ZAR1,include.mean=FALSE,trace=F)
> res4=residuals(m10,standardize=TRUE)
> pairs( res+res1+res2+res3+res4)

```

Multivariate elliptical stable parameter estimation

```

> br<-matrix(c(res,res1),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(br, method1d=1)
> bi<-matrix(c(res,res2),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(bi, method1d=1)
> bc<-matrix(c(res,res3),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(bc, method1d=1)
> bz<-matrix(c(res,res4),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(bz, method1d=1)
> ri<-matrix(c(res1,res2),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(ri, method1d=1)
> rc<-matrix(c(res1,res3),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(rc, method1d=1)
> rz<-matrix(c(res1,res4),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(rz, method1d=1)
> ic<-matrix(c(res2,res3),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(ic, method1d=1)
> iz<-matrix(c(res2,res4),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(iz, method1d=1)
> cz<-matrix(c(res3,res4),nrow=2,ncol = 1324)
> mvstable.fit.elliptical(cz, method1d=1)

```

Estimated density and contour plots

```
>stable.test(xx <- seq(-0.04, 0.04, 0.002)
d.ell <- mvstable.elliptical(alpha =1.871113, R = matrix(c(0.4168822 ,0.0539122,0.0539122,
0.3965146), 2, 2), delta = c( 0.01444671,0.07260969))
pdf.surface.plots(d.ell, xx, "elliptical"))
```